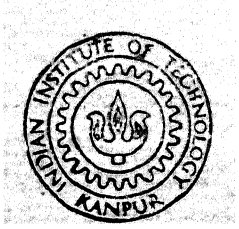
NUMERICAL TREATMENT OF SINGULARLY PERTURBED TWO - POINT BOUNDARY - VALUE SYSTEMS

by

Arindama Singh



DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY KANPUR

DECEMBER, 1988

NUMERICAL TREATMENT OF SINGULARLY PERTURBED TWO - POINT BOUNDARY - VALUE SYSTEMS

A Thesis Submitted
In Partial Fulfilment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

by

Arindama Singh

to the

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY KANPUR
DECEMBER, 1988

CERTIFICATE

Certified that the thesis entitled "Numerical Treatment of Singularly Perturbed Two-Point Boundary-Value Systems" by Arindama Singh has been carried under my supervision and has not been submitted elsewhere for a degree or diploma.

December 1988

Mohan K. Kadalbajoo
Department of Mathematics
Indian Institute of Technology
Kanpur - 208016
India

ACKNOWLEDGEMENTS

While responsibility for any errors or shortcomings in this thesis rests with me, its inspiration and content owe much to my thesis supervisor Dr. Mohan K. Kadalbajoo. I wish to express my indebtedness to Prof. P.C. Das for his help and guidance during Dr. Kadalbajoo's absence. Thanks are due to Mr. K.N. Islam for his excellent typing.

A.S.

List of Publications From This Thesis

1. A Cutting Point Technique for Solving Nonlinear State Regulator Problems.
   - IMA J. Mathematical Control and Information, (1987), Vol. 44, PP. 183 - 194.

2. A Boundary Value Technique for Solving Singularly Perturbed Fixed End-Point Optimal Control Problems.
   - Optimal Control - Applications and Methods,
   - Accepted for Publication (Vol. 9).

3. A Boundary Value Technique to Solve Linear State Regulator Problems.
   - J. Optimization Theory and Application,
   - Accepted for Publication.

4. A Numerical Method for Solving Singularly Perturbed Systems of Nonlinear Two-Point Boundary-Value Problems.
   - Communicated.

5. Partial Decoupling of Slow and Fast Variables.
   - Communicated.

6. An Estimate of Boundary Layer Thickness for Linear Singularly Perturbed Two-Point Boundary-Value Problems.
   - Communicated.

TABLE OF CONTENTS

List of Tables

List of Figures

Synopsis

of the thesis entitled

NUMERICAL TREATMENT OF SINGULARLY PERTURBED TWO-POINT BOUNDARY-VALUE SYSTEMS

by

Arindama Singh
Department of Mathematics
Indian Institute of Technology, Kanpur-208016, India

Perhaps the human beings form a distinguished species of the genus 'Animals', because they possess the power of abstraction. It is this power of abstraction that induces them to prefer mathematical models to describe real life problems. Then they analyse the models to understand and predict real life situations.

Very often it happens that the models contain small parameters which one would otherwise like to neglect, (i.e., replacing small parameters by zero). It is all well if one is justified to do so. What if one is not justified in doing so? This is the beginning of singular perturbation phenomenon.

Assume, for simplicity, that P_ε is a problem expressed by ordinary differential equations, subject

to appropriate boundary conditions, where only one small parameter ε is involved. P_ε is a singular perturbation problem, if the solution S_ε of P_ε does not converge (uniformly) to the solution S_o of the reduced problem P_o obtained from P_ε by letting ε approach zero. However, in some cases, corrections to S_o could be done so that it approximates S_ε within, say, $O(\varepsilon^N)$ accuracy. Asymptotic series solutions and matched asymptotic expansion techniques were found to be somewhat accurate in this direction. But the difficulties in using such techniques are well known; they lack simplicity and are very demanding on the part of the user.

Numerical analysis accepted this new challenge. Consequently, numerical methods have been devised for special classes of singular perturbation problems. Beyond these special classes of problems most of these methods do not work. Methods which fulfil this need of transcending the subclasses of problems are neither simple nor cheap.

In this thesis a general purpose method for solving systems of singularly perturbed two-point boundary-value problems in ordinary differential equations is proposed, analyzed and illustrated. Methods are also prescribed for some subclasses of

this problem class, keeping the above requirements implicit. The thesis comprises seven chapters.

Chapter 1 of this thesis presents motivation, a brief survey of the asymptotic and numerical analysis of such problems and summary of the present work. Results from various sources that are used in later chapters are also compiled to facilitate understanding the materials of this thesis.

In Chapter 2, a simple partial decoupling transformation is constructed which decouples the slowly varying component from the original linear singularly perturbed system. Euler's Implicit (forward) and Explicit (backward) methods are shown to be accurate (of $O(h)$ or $O(h^2)$) for the rapidly varying component. Non-commutativity of the difference schemes and the decoupling transformations is also discussed.

Based on the results of Chapter 2, an algorithm for singularly perturbed linear systems of two-point boundary-value problems is presented in Chapter 3. The effectiveness of the algorithm is demonstrated by some illustrative examples.

In Chapter 4, two special methods, namely, Boundary Value Technique and Cutting Point Technique are proposed. These techniques are applicable when

the problems considered involve no turning points. These methods exploit the boundary-layer thickness, an indirect estimate of which is also derived in this chapter.

In Chapter 5, the methods of preceding chapters are applied to non-linear singularly perturbed two-point boundary-value systems via a modified version of Newton's iteration. A model example of non-linear problem is solved to demonstrate the effectiveness of the methods.

In Chapter 6, these methods are applied to optimal control problems both linear and non-linear.

Chapter 7 provides a conclusion to this thesis and indicates scopes for further research.

It is observed that the numerical methods presented have been found to be efficient over the conventional methods. The methods are at the same time, conceptually simple and need a very modest amount of problem preparation for computational purposes. Model problems of linear and non-linear types have been solved and numerical results are presented wherever necessary. The predicted accuracy could always be achieved without much computational effort.

All the numerical results presented in this thesis have been computed on DEC-1090 Computer system available at the Indian Institute of Technology, Kanpur, India.

Chapter 1

INTRODUCTION

1.1 Introduction

It is a fancy of pure mathematicians to invent and do such mathematics which has no obvious applications: mathematics for the sake of mathematics. However, most of mathematics that we do today have its origin in real life problems, or say, applications. Singular perturbation came into existence, not by fancy, but by necessity, thereby it belongs to the latter category. Ever since Prandtl's [96] work in the beginning of this century, singular perturbation techniques have been a traditional tool of fluid dynamics. These techniques entered into various other areas of application, where ofcourse, the same terminology of 'boundary layer', 'interior layer', 'outer' and 'inner' expansions were already in use. Here, just to arouse one's curiosity, we might mention its use in various applied areas such as fluid dynamics, plasticity, chemical reactor theory, nuclear reactor theory, plasma physics, aerodynamics, meteorology, rarefied gas dynamics, diffraction theory, reaction-diffusion process, non-equilibrium and radiating flows and etc. Later, we will consider an application area of singular perturbation : optimal control theory. Singular perturbation

problems that cropped up in these and many other areas were treated almost exhaustively by asymptotic analysis. A breakthrough occured when numerical analysts turned their attention towards these problems. It was a challenge to numerical analysis. (We will shortly see why.). Unlike the first expectation to see these two approaches as water-tight compartments, it turned out that wherever they are simultaneously applied, they rather stand complementary to each other. This combined effort turned out to be an efficient mathematical aparatus in investigating singular perturbation phenomenon in various areas of engineering.

The purpose of this chapter is to introduce and describe the problems those are treated in this thesis. Also a brief survey of the literature is presented. It should not only point out some of the problems that were or are being treated but also the manner in which these are being treated. Only a summary of some recent methods is presented. Obviously, a selection of techniques is done implicitly to present this summary, which is in some way relevant to the work pursued in this thesis. The methods so far developed, both prescribe as well as proscribe the approaches that should or should not be adopted, by recognizing the problem classes. The problems at hand are such that, even an unskilled eye would see their worth as to be paid such

a singular attention. We present here two examples to justify the singular attention and having a peep into the main difficulties involved. Without much ado, let us hear the monologues of first two explanatory characters of our drama.

Example 1. ([32])

Consider the following system of two-point boundary value problem :

$$\dot{y}_1 = y_1 + y_2 \qquad (1.1a)$$

$$0 < t < 1$$

$$\varepsilon\dot{y}_2 = - y_2 \qquad (1.1b)$$

with boundary conditions

$$y_1(0) = \alpha , \qquad y_2(1) = \beta \qquad (1.2)$$

where '.' denotes d/dt, α, β are some fixed real numbers, and ε is a small positive parameter $(0 < \varepsilon << 1)$.

Before going to any of the numerical methods, it is advisable to observe the equations carefully. This could best be done if we could get the exact solution. Fortunately, we do have the exact solution, i.e.,

$$y_1(t,\varepsilon) = (\alpha + \beta\, \varepsilon \exp(1/\varepsilon)/(1+\varepsilon))\, \exp(t) - \beta\, \varepsilon\, \exp((1-t)/\varepsilon)/(1+\varepsilon) \qquad (1.3a)$$

$$y_2(t,\varepsilon) = \beta\, \exp((1-t)/\varepsilon) \qquad (1.3b)$$

Now it is clear that y_1 changes very rapidly everywhere on $(0,1]$ whereas y_2 blows up only in a small neighbourhood of $t = 1$ depending upon how small ε is. This is a typical singular perturbation phenomenon, though the name 'Singular Perturbation' was given due to some other reason (cf. §1.2). However, for numerical methods to be efficient, at least to solve such problems accurately, it is a warning that standard two-point boundary-value techniques with uniform grid should never be used, unless the grid length is very small (then ofcourse, the round-off errors could surpass the accuracy level); and in all cases, the 'small neighbourhood' where y_2 blows up should receive special care. Observing (1.1) through a quantitative eye, we find that the failure of the standard techniques can be attributed to the difference in order of magnitudes of the eigenvalues of the system matrix. The eigenvalues are 1 and $-1/\varepsilon$; when $\varepsilon \ll 1$, this difference in order of magnitudes will lead to numerical instability of the standard methods. In fact, this is one of the main difficulties involved in such problems. Intentionally, to give a chill in one's spine, we have chosen such a simple example, where the system matrix is time-independent. In the next example, we take the system matrix to be time-dependent and enter into a bit dirty (but not unaesthetic) water.

Example 2. ([58])

Consider the following second order two-point boundary-value problem :

$$\varepsilon \ddot{y} + t\dot{y} - y = 0 \qquad -1 < t < 1 \tag{1.4}$$

$$y(-1) = 1,\ y(1) = 2 \tag{1.5}$$

The exact solution of (1.4-1.5) is :

$$y(t,\varepsilon) = c_1 t + c_2 (\exp(-t^2/(2\varepsilon)) + (t/\varepsilon) \int_{-1}^{t} \exp(-s^2/(2\varepsilon))ds) \tag{1.6}$$

where $c_1 = -1 + 3\exp(-1/(2\varepsilon))/A$ (1.7)

$$c_2 = 3/A \tag{1.8}$$

$$A = 2\exp(-1/(2\varepsilon)) + I(\varepsilon)/\varepsilon \tag{1.9}$$

with $I(\varepsilon) = \int_{-1}^{1} \exp(-s^2/(2\varepsilon))ds$ (1.10)

Note that $I(\varepsilon) = \sqrt{2\pi\varepsilon}$ + T.S.T. (1.11)

where T.S.T. is an abbreviation for 'transcendentally small terms'. With the help of (1.11) the solution $y(t,\varepsilon)$ can be simplified as

$$y = \begin{cases} -t - \text{T.S.T.} & t < 0 \\ 3\sqrt{\varepsilon/(2\pi)} + \text{T.S.T.} & t = 0 \\ -2t + \text{T.S.T.} & t > 0 \end{cases} \tag{1.12}$$

This describes a 'corner' at t = 0. Here again, we note that the eigenvalues of the system matrix are no longer time-independent as in Example 1. In fact, as t varies, the eigenvalues change their orders of magnitude on the interval of integration [-1,1]. This is another main difficulty in the presence of which numerical stability of the standard techniques is jeopardised.

The troublesome region in Example 1 is situated around the point t = 1, called a boundary layer. The situation in Example 2 is termed, in general, as a turning point (t = 0), which gives rise to interior or corner layers asymptotically.

Naturally enough, the failure of standard techniques forces us to ask for better (specialized) techniques using non-uniform grids. We should not also forget to see from a user's point of view. A user might not entertain himself to see the troublesome regions beforehand. So we rather ask for a method or methods that indeed, succeed by finding out these troublesome regions simultaneously. This question will be answered in the following sections.

There was a time when asymptotics had the monopoly in tackling such problems, and one had to seek an asymptotic solution unless one had already found out an exact (analytic) solution. So it is not advisable to

neglect asymptotics altogether. Since our main emphasis is on numerical treatment, we will pass asymptotics over with a rapid pace. And in the later sections we will come across some of the numerical approaches that have been taken to tackle such difficult problems.

1.2 Asymptotic Treatment of Singular Perturbation

Let P_ε be a problem in some differential model, which depends upon a positive small parameter ε $(0 < \varepsilon << 1)$. Since ε is small compared to 1, it is natural to consider also the degenerate or reduced problem P_0, obtained from P_ε by replacing every occurance of ε in P_ε with zero; here we assume that the problem P_0 so obtained is meaningful. One then asks whether the solution S_ε of P_ε does converge uniformly to the (reduced) solution S_0 of the reduced problem P_0, as $\varepsilon \to 0$. If the answer is in the affirmative, we call P_ε, a regular perturbation problem, whereas, if the answer is in the negative, we call P_ε a nonregular or singular perturbation problem. The word 'perturbation' is used viewing P_0 as the original or unperturbed problem and thereby P_ε, its perturbation.

In the following we will be concerned with the study of the latter, namely, the singular perturbation. In particular we will restrict our attention to this singular perturbation phenomenon in ordinary differential

equations and that to, only the two-point boundary-value problems (TPBVP). Unless and otherwise stated explicitly, singular perturbation problem will mean singularly perturbed two-point boundary-value problem.

To be specific, we consider the problems of the type

$$\dot{x} = f(x, y, t, \varepsilon) \qquad (2.1a)$$

$$a < t < b.$$

$$\varepsilon\dot{y} = g(x, y, t, \varepsilon) \qquad (2.1b)$$

with boundary conditions

$$F(x(a), x(b), y(a), y(b), \varepsilon) = 0 \qquad (2.2)$$

where $t \in \mathbb{R}$ the real numbers, represents time, $x(t,\varepsilon)$ and $y(t,\varepsilon)$ are m and n dimensional real vector functions respectively; '.' denotes d/dt; f and g are smooth real vector functions linear or non-linear of dimensions m and n respectively, and F in (2.2) represents the boundary conditions so that (2.1) might be solved uniquely. In case, F is linear, it should represent (m+n) - linearly independent boundary conditions.

In an excellent review article by O'Malley [85], history of asymptotic methods applied to such problems is presented. Starting from Prandtl's [96] paper on fluid dynamical boundary layers, which in fact, began

the mathematical study of such problems, O'Malley lists many other important works on asymptotic methods. In asymptotics, a series solution of the form

$$z(t,\varepsilon) = \sum_{k=o}^{\infty} z_k(t).\varepsilon^k \ , \quad z = \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2.3)$$

is tried and under certain hypotheses such a series is found out for a specific member of the class of problems (2.1 - 2.2). The solution can be found out, for practical purposes, upto $O(\varepsilon^L)$ accuracy, for some finite L (typically $L = 1$ or 2). This rather concentrates on the form of the solution, that too in a series form as in (2.3). However, since practically we truncate the series in (2.3) after ε or ε^2, the terms left are represented as $O(\varepsilon^L)$. This rather concentrates on the nature of the solution with respect to ε .

Kevorkian and Cole [58], for example emphasized on the nature of the solution on the whole interval $[a,b]$, thereby studying the solutions from both the angles namely t and ε . Sometimes, it happens that a series of the type

$$z(t,\varepsilon) = \sum_{k=o}^{\infty} z_k(t).(\sqrt{\varepsilon})^k \ , \quad z = \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2.4)$$

approximates solution better than that in (2.3). Pictorially, this represents the rapid change of the solution (see Example 1, § 1.1) in a stretched independent

variable, which stretches t by a power of ε. In (2.4) this power of ε is taken to be $\sqrt{\varepsilon}$, the stretching parameter. In a series solution given by (2.3), the stretching parameter is ε itself. This idea can also be found in Eckhaus [27,28], though with more rigor.

A simplified matching procedure has been presented by Van Dyke [107] to tackle these problems efficiently. Now, in application, this matched asymptotic expansion process is used widely for such problems. Also, a number of books, monographs and research reports have appeared which handle these problems mathematically. In a review article by Kokotovic [60], one can find an exhaustive bibliography on this topic. For an applied man new to this area, Ardema [4] provides a good introduction. Before going to matching, it is worthwhile to look at Example 1 of Section 1.1 once again. The second component of the solution vector, i.e., $y_2(t,\varepsilon)$ blows up in a boundary layer at $t = 1$. Suppose we scale the equations (stretch the boundary layer) by a new independent variable $\sigma = (1-t)/\varepsilon$. Then $y_2(t,\varepsilon)$ can be represented by $y_2(\sigma,\varepsilon) = \beta \exp(\sigma)$ which does not blow up in σ. Note that a scaling in equations produces also a scaling in the solutions. This region, the boundary layer region, is also referred to as the inner region;

and the other region, beyond the boundary layers is termed as the outer region. In general, it is striking to take the stretched problem and solve it in the boundary layer region. This procedure is followed when the technique of matched asymptotic expansions is used. In the inner region we have the stretched problem and in the outer region, the original one. But number of boundary conditions is still the same. To solve these problems, one extra boundary condition is imposed, which comes from the matching of two solutions of these two problems. The constants of integration are determined by matching these solutions in an 'overlap region', i.e., where both the stretched and the unstretched solutions match. This matching is necessary since nothing about the exactness of the regions is known beforehand. These names such as inner region and outer region expansions, overlap regions etc., in fact, came from various application areas like fluid dynamics, optimal control, nuclear reactor theory etc.

The results obtained through matching coincide with those known by intuitive folkways of various applications. The most referred to texts include, for example, Bellman [13], Nayfeh [76,77], O'Malley [82], Eckhaus [27,28], Vasileva [108], Tikhnov [105], Erdelyi [30], Ferguson [31].

The main problem with asymptotic techniques is that it is very demanding on the part of the user. The matching procedure is still not very simple to use. As Eckhaus [28] pointed out, one should know beforehand which stretching factor (some power of ε, i.e., ε^r, r: a rational number) is to be used. Our aim is to present numerical methods for users. For the fact that we do not miss a great name, this section is included. For more details regarding asymptotic expansion techniques, one should go through many of the literature not cited here. The key to those might be found out from O'Malley [85] and Kokotovic [60].

1.3 Numerical Treatment of Singularly Perturbed Second Order TPBVPs.

Leaving aside the analytic approaches or qualitative theories that have been applied to singularly perturbed two-point boundary value problems, we are forced to consider the asymptotics and the numerical means. Recently,Non-standard Analysis has been applied to singular perturbations [14,73]; but this only informs us about the geometrical behaviour of the solutions of a specific class of problems. For practical purposes, one needs quantitative means of solving the problems. In asymptotic analysis of

such problems, the solution is represented as an additive sum of a certain (quantitative) function, and a term , the remainder or the residual term, which represents more of a property than a quantity [106, Ch-VII], like $O(\varepsilon^2)$ and etc. Numerical analysis starts when we need to know the solution quantitatively and accurately. More precisely, while Numerical analysis provides us information regarding the solutions through quantitative - coloured glasses, Asymptotic analysis projects the solutions on a semi-quantitative, semi-qualitative tinted screen. However, numerical methods are intended at a broader class of problems and at the same time demands less from the user. Again, when for a specific problem one finds it very difficult to ride through asymptotics, it is natural to expect help from Numerical analysis.

Unfortunately and not unexpectedly, the numerical analysis of singular perturbation is not that easy. As was remarked earlier, a non-uniform mesh is needed for the numerical schemes to produce accurate approximations. Note that this idea, if carried out might also reduce the number of grid-points, which is more or less necessary for an effective computation. Many such grid-generation processes have been used. Pearson [92], for example, used

variable grid-length combined with a three-point difference scheme for the problems of the type

$$\varepsilon \ddot{x} + a(t)\dot{x} + b(t)x = f(t) \qquad -1 < t < 1 \qquad (3.1a)$$

$$x(-1) = A \ , \quad x(1) = B \qquad (3.1b)$$

where '.' denotes d/dt and ε is a small positive parameter. Pearson's procedure in constructing a variable grid is conceptually simple. We require the grid to be dense in those regions where the solution changes rapidly. So the first step would be to know where does it really happen. For this purpose, he takes a uniform grid (with a fixed and modest ε) and solves the problem numerically with the help of a three-point difference-scheme. To test the rapid change in the solution, a tolerance limit δ is prescribed for the difference $d_i = |x_{i+1} - x_i|$. $d_i > \delta$ would mean refining the grid by introducing grid points between t_i and t_{i+1}. A smoothing process is also carried out to preclude abrupt changes within any mesh-interval. Then for smaller values of ε, one takes the previous grid instead of a new uniform grid and then repeat the whole procedure once again. Though this method is very simple to use, the major drawback is cost, i.e., in terms of machine-time. Pearson [93] also used this method for non-linear problems of the type

$$F(t, x, \dot{x}, \varepsilon\ddot{x}) = 0 \qquad 0 < t < 1 \qquad (3.2a)$$

$$x(0) = A, \quad x(1) = B \qquad (3.2b)$$

The number of grid-points necessary for the numerical solution of these problems through this method is very large. For example, to solve a simple problem

$$10^{-6}\ddot{x} + (\dot{x})^2 = 1 \qquad 0 < t < 1 \qquad (3.3a)$$

$$x(0) = x(1) = 1 \qquad (3.3b)$$

upto five accurate digits, he had to construct a grid consisting of 4000 points.

It may be remarked that a uniform grid is not expected to work satisfactorily due to the numerical instability of the problems considered. However, a fresh look at the numerical instability and an inertia towards using uniform mesh leads one to consider a directional difference scheme rather than the centered difference schemes [24].

This idea was developed further by Dorr et al. [26]. They applied a discrete maximum principle to obtain estimates on the numerical solutions of the equations (3.1). The choice of the direction in using the directional difference schemes (i.e., backward or forward) depends upon the sign of a(t) in (3.1). This also establishes the link with the asymptotic theory, as it is known that a boundary layer occurs

at the left or right end point of the interval according as a(t) is negative or positive throughout the interval of integration [0,1]. The main drawback of this approach is that it can only be applied to a very restricted subclass of the classes of problems (3.1-3.2), particularly, where the maximum principle works. In the other direction, Dorr et al. [26] extended their results for the problem class which admits turning points and also to quasi-linear problems.

Researchers sometimes feel reluctant to treat boundary-value problems and in particular, two-point boundary-value problems, because, in most cases, these can be transformed to initial value problems by the so called shooting methods. For example, Cohen [19], Dorr and Parter [25] have considered the application of shooting methods for non-linear problems of the type

$$\varepsilon\ddot{x} + f(t, x, \dot{x})\dot{x} = 0 \qquad (3.4)$$

and coupled equations

$$\dot{x} = f(t, x, y) \qquad (3.5a)$$

$$\varepsilon\dot{y} + g(t, x, \dot{x})\dot{y} = c(t, x, \dot{x})y \qquad (3.5b)$$

with appropriate boundary conditions. Kopell and

Parter [64] analysed the problem

$$\varepsilon\ddot{x} = (x^2 - t^2)\,\dot{x} \qquad (3.6a)$$

$$x(-1) = A, \quad x(1) = B \qquad (3.6b)$$

again applying shooting methods.

Numerical methods have also been devised which first approximate the co-efficients $a(t)$, $b(t)$ in (3.1) by polynomials and then apply some difference scheme. Ortiz [90], for example, estimated the errors in applying Tau method to the class of singular perturbation problems

$$\ddot{x} + \varepsilon^{-1}p(t)\,x = 0 \qquad 0 < t < 1 \qquad (3.7a)$$

$$x(0) = 1, \quad x(1) = 0 \qquad (3.7b)$$

where $0 < \varepsilon << 1$ and $p(t)$ is a polynomial approximation of a function not identically equal to zero on $[0,1]$.

In asymptotics, the reduced problem plays an essential role. The solution of the reduced problem is called, as one expects, the reduced solution. It happens that in the regions beyond the layers, the reduced solution represents or approximates the original solution fairly well. Flaherty and O'Malley [35] have given a numerical algorithm for singularly perturbed second order TPBVPs, or

more generally, for second order stiff TPBVPs, based on the reduced problem. The characteristic roots of the reduced problem enables them to obtain the boundary layer correction terms numerically. These correction terms are corrections to the reduced solution(s) in respective regions. The reduced solution could also be obtained from the reduced problem numerically. This is a beautiful example of the cross-breeding of the two approaches : asymptotic and numerical.

Carried on by the asymptotic experiences, Lorenz [72] has considered the numerical solution of the problem

$$\varepsilon\ddot{x} = \alpha(t,x)\,\dot{x} + \beta(t,x) = 0 \qquad 0 < t < 1 \qquad (3.8a)$$

$$x(0) = A, \quad x(1) = B \qquad (3.8b)$$

He splits the interval of integration [0,1] into two parts and then on one part solves the reduced problem, whereas on the other part, solves the original problem. The latter part being stretched appropriately before the difference schemes are applied.

Recently, Kadalbajoo and Reddy [47-50] have given some methods for singularly perturbed two-point boundary-value problems of second order, which go in the same vein of observing the problem from both the asymptotic as well as numerical points of view.

Also many non-asymptotic methods have been invented. For example, Roberts [97] has given a method, there called, Boundary Value Technique for singularly perturbed TPBVP's. He has also modified the method a bit so that it could be applied to problems of an extended class [98-99]. These methods have been found successful even in many non-linear cases.

Among others, the three-point difference schemes of Kellogg et al. [57], the chopping procedures of Sakai [100], Sakai et al. [101], exponential tri-diagonal difference scheme of Berger et al. [15], collocation methods using cubic polynomials and cubic splines by Flaherty and Mathon [34], box and trapezoidal schemes studied by Weiss [113] note-worthy. Axelsson and Carey [12] constructed a regular modified problem to discuss the boundary layers separately and derived thereby the estimates for boundary layers which are uniform in terms of the small (but immensely notorious) parameter ε .

1.4 Numerical Treatment of Systems of Singularly Perturbed TPBVPs.

Compared to numerical treatment for handling systems of singularly perturbed two-point boundary

value problems much has been done by the asymptotics. In various treatments of such problems, several problem classes have been considered. Among these problem classes, much attention has been paid to linear homogeneous systems because of their occurence in a wide range of applications. For example, in Optimal Control Theory, when a scalar (quadratic) cost is optimized subject to the states modelled by ordinary differential equations containing a small parameter ε $(0 < \varepsilon << 1)$, which typically represents small mass, inductance, capacitance etc., it happens that the resulting differential equations to be solved become singularly perturbed. Kokotovic [60] claims that almost all control problems are singularly perturbed; one does not recognize them because it is customary to neglect the small parameters. Though several such parameters might appear in a specific model giving rise to singularly perturbed systems with several parameters, we will be interested here in describing only one-parameter cases.

The linear homogeneous system of singularly perturbed TPBVP's can be represented as follows :

$$\dot{x} = Ax + By \qquad 0 < t < 1 \tag{4.1a}$$

$$\varepsilon\dot{y} = Cx + Dy \tag{4.1b}$$

$$B_1 \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} + B_2 \begin{bmatrix} x(1) \\ y(1) \end{bmatrix} = B_3 \tag{4.1c}$$

where $x(t,\varepsilon)$, $y(t,\varepsilon)$ are vector functions of dimensions m,n ; $A(t,\varepsilon)$, $B(t,\varepsilon)$, $C(t,\varepsilon)$, $D(t,\varepsilon)$ are given matrix functions of dimensions mxm, mxn, nxm, nxn; and (4.1c) represents (m+n) - linearly independent boundary conditions; '.' denotes d/dt, and ε is a small positive parameter, $0 < \varepsilon << 1$.

Asymptotic analysis of this problem (4.1) reveals the fact that if, A, B, C, D are smooth bounded matrices, then $x(t,\varepsilon)$ is a slowly varying component of the solution and $y(t,\varepsilon)$ is the rapidly (fast) varying component. Under certain hypotheses the asymptotic solution has been obtained. To name a few, credits might go to Vasileva [108], Tikhonov [105], (see also [111]), O'Malley [82], Kokotovic et al. [62], Hopensteadt [44], Sannuti [103], Finden [33], Nayfeh [76], Nipp [79] and for their other works in the related fields as well as that of their co-workers. However, for a detailed bibliography one can see Kokotovic [60].

Various assumptions have been taken in order to ensure that the solution $z_\varepsilon(t)$ of (4.1) approaches the reduced solution $z_o(t)$ of the corresponding reduced system $\left(z = \begin{bmatrix} x \\ y \end{bmatrix}\right)$

$$\dot{x} = Ax + By \qquad (4.2a)$$

$$0 = Cx + Dy \qquad (4.2b)$$

with appropriate choice of m-boundary conditions from

(4.1c), in the limit as $\varepsilon \longrightarrow 0+$. Recently, Campbell [17] has given a new approach for initial value problems. Also, differential inequalities have been applied to guarantee the existence/uniqueness as well as this limiting behaviour of the solution $z_\varepsilon(t)$.

Howes [45] considers singularly perturbed second order non-linear systems of the type

$$\varepsilon \ddot{x} + F(x)\,\dot{x} + g(t,x) = 0 \qquad a < t < b \qquad (4.3)$$

where F is a matrix valued function. Sufficient conditions are given in [45] in order that a solution of (4.3) displays a layer behaviour, i.e., non-uniform dependence of the solution on t and ε as $\varepsilon \longrightarrow 0+$.

Flaherty and O'Malley [35] combined the asymptotic analysis and numerical analysis, in the sense that, to construct a numerical solution, they used the asymptotic information. Though, they have generalized their method to quasi-linear scalar case, they have not extended it to systems. However, as we have remarked earlier, the asymptotic analysis of singularly perturbed systems of two-point boundary-value problems demand, also, a variable mesh construction for the effectiveness of the numerical algorithms. This means, in different regions of the interval of integration, different grid-lengths

must be used as well as a stretching of the small regions helps in achieving a comparatively flawless computation. Abrahamson et al. [1], for example, analysed the second order system

$$\varepsilon \ddot{x} + A(t)\,\dot{x} + B(t)\,x = F(t) \qquad 0 < t < 1 \tag{4.4a}$$

$$x(0) = \alpha, \qquad x(1) = \beta. \tag{4.4b}$$

They assume that

$$A(t) = \operatorname{diag}(A^{I}(t),\ A^{II}(t)) \tag{4.5a}$$

$$A = A^{*} \tag{4.5b}$$

where A^{*} is the adjoint of A, and there is some constant $\eta > 0$ such that

$$A^{I} \leq \eta < 0 \qquad \text{and} \qquad A^{II} \geq \eta > 0 \tag{4.5c}$$

for every t, $0 \leq t \leq 1$; where e.g. $A^{I} \leq \eta$ means that all the entries in A^{I} are less than or equal to η.

Introduce the notation

$$x = \begin{bmatrix} x^{I} \\ x^{II} \end{bmatrix}, \qquad F = \begin{bmatrix} F^{I} \\ F^{II} \end{bmatrix}$$

corresponding to the partition of A. This partition plays its role in letting the boundary layers appear in narrow regions at $x = 0$ and at $x = 1$ for the

variables x^{II} and x^{I} respectively. This has been proved by finding an a priori estimate of the solution of (4.4). In constructing numerical algorithms it is important that the numerical algorithm should be able to work efficiently when the mesh size $h >> \varepsilon$. Otherwise, one has to take h very small to carry out the computation thereby causing roundoff errors and resulting in prohibitive computer time.

Abrahamson et al. [1] used the following difference approximation. The problem (4.4) is approximated by

$$(\varepsilon + h\, C_k) D_+ D_- u_k + A_k D_o u_k + L_{1k}(B,u) = L_{2k}(F) \qquad (4.6a)$$

$$u_o = \alpha, \qquad u_N = \beta \qquad (4.6b)$$

where $C(t) = \mathrm{diag}(C^{I}(t), C^{II}(t))$ is a positive definite matrix such that $C(t) \geq \frac{1}{2}\mathrm{diag}(-A^{I}(t), A^{II}(t))$, D_+, D_- and D_o are the usual forward, backward and central difference operators and, L_1, L_2 are linear operators of the form

$$L_{1k}(B,u) = B_{-1k}\, u_{k-1} + B_{ok}\, u_k + B_{1k}\, u_{k+1}$$

$$L_{2k}(F) = C_{-1}\, F(t_k - \tfrac{h}{2}) + C_o\, F(t_k) + C_1 F(t_k + h/2).$$

The coefficients B_{ij}, C_i's are chosen to make (4.6) as accurate as possible; it would also make the boundary layers as sharp as possible. They have used Richardson extrapolation in the interior of the interval to achieve the accuracy upto $O(\varepsilon)$. The smoothness of the numerical solution has also been shown except in the boundary layer regions of thickness of $O(\varepsilon|\ln\varepsilon|)$ and the accuracy achieved is $O(\varepsilon)$, and the rate of convergence is $O(h)$.

Following this approach as well as Pearson's [92-93] approach of variable mesh construction, Kreiss and Kreiss [67] developed a method for this class of problems. A variable mesh is suggested instead of fixing the mesh beforehand. Starting with some reference mesh, a difference scheme of the type (4.6) is used to compute a numerical solution. Then, a new mesh is constructed by adding or deleting mesh points according to the variation of the numerical solution. This novel feature enables them to extend the method to the cases when the condition $A = A^*$ in (4.5b) is not satisfied; as well as this method is also applicable to the non-homogeneous form corresponding to (4.1). The assumptions (4.5) are connected with the two main difficulties associated with systems like (4.1) and its non-homogeneous form. The

difficulties encountered are (see the Examples 1, 2 in § 1.1) : first, the eigenvalues of the system matrix

$$\begin{bmatrix} A & B \\ \varepsilon^{-1}C & \varepsilon^{-1}D \end{bmatrix}$$

vary widely in magnitude, and secondly, these eigenvalues change their orders of magnitude when t varies along [0,1] (see [66, 75]).

Methods have also been developed for a special class of problems, namely, singular-singularly perturbed two-point boundary value systems. For example, O'Malley [84] considers systems of the type

$$\dot{x} = A(t,\varepsilon)x + B(t,\varepsilon)y + C(t,\varepsilon) \qquad (4.7a)$$

$$\varepsilon\dot{y} = E(t,\varepsilon)x + F(t,\varepsilon)y + G(t,\varepsilon) \qquad (4.7b)$$

(4.7) is called singular by assuming that the matrix $F(t,0)$ is singular. A full set of asymptotic solutions is constructed assuming further that $F(t,0)$ is block-diagonalizable and the reduced problem is consistent and a stability condition (defined in [84]) holds. This would give rise to numerical methods mentioned earlier, for example, as in [35].

Ascher [7] considers the following system of TPBVP's :

$$\varepsilon \dot{x} = A(t,\varepsilon)x + f(t,\varepsilon) \qquad 0 < t < 1 \qquad (4.8a)$$

$$B_0(\varepsilon)\, x(0,\varepsilon) + B_1(\varepsilon)\, x(1,\varepsilon) = \beta(\varepsilon) \qquad (4.8b)$$

where $x(t,\varepsilon)$ is a vector valued function, and $f(t,\varepsilon)$ is a vector valued linear or non-linear function, $A(t,\varepsilon)$, a matrix valued function of appropriate dimensions; and (4.8b) represents boundary conditions. The reduced problem corresponding to (4.8) is assumed to be singular. For numerical approximation, families of symmetric difference schemes are used. Convergence results for the regular singularly perturbed case are extended with reference to a modification in mesh selection. Convergence of mid-point rule and trapezoidal rule are proved. Ascher proved this convergence by assuming some conditions on the difference schemes rather than on the problem (4.8) itself.

Recently Kreiss et al. [68] have given a method for stiff two-point boundary-value problems. This method uses a grid-generation process closely related to that of Abrahamson et al. [1]. The difference scheme is a combination of Euler's Implicit, Euler's

Explicit and Trapezoidal rules. However, the grid-generation process consumes too much of machine time and it also involves a large number of function-evaluations.

The non-linear singularly perturbed two-point boundary-value systems are treated with shooting methods, as is the case for non-linear stiff counter-parts. Using Newton's Iteration [89], the non-linear problems could be realized as the limit of a sequence of linear problems and then the methods for linear problems could be used [68].

Due to the undoubted status of shooting, mathe-maticians are reluctant to accept the adjective 'stiff' being applied to two-point boundary-value problems (§ 8.2 in [2]). However, the case of singular perturbation is non-controversial.

Among others, difference schemes of Dorr [24-25], spline collocation codes of Ascher et al. [6-7, 11] could very well be incorporated for systems of general linear or non-linear singularly perturbed TPBVP's. In fact, the collocation codes were already constructed for this problem class, see e.g., Ascher and Weiss [8-10]. Centred difference schemes were analyzed by Kreiss [65] when applied to singular perturbations.

An inclusion relationship between perturbed collocation and Runge-Kutta methods was investigated by Nørsett and Wanner [80]. Uniform stability of discrete and continuous singularly perturbed boundary-value problems was studied by Niederdrenk and Yserentant [78]. Weiss [113] analyzed the box scheme and the trapezoidal rule for linear singularly perturbed boundary-value problems. Smith [104] considered a singularly perturbed boundary-value problem arising in the physical theory of semiconductors. For difficult topics like singular-singularly perturbed TPBVP's, Vasileva and Butuzov [110] have given the asymptotic solutions. The one-sided difference schemes are shown to be applicable in nonlinear singularly perturbed problems by Osher [91].

1.5 Summary of the Thesis

In the previous sections, we have come across various classes of problems as well as asymptotic and numerical methods for their solutions. Both the asymptotic and numerical methods aim at finding accurate solutions to concerned problem (classes). The accuracy in asymptotic expansions is achieved upto an order in some power of ε $(O(\varepsilon^n))$ and in numerical techniques, the accuracy is achieved upto an

order in some power of the maximum grid length h ($O(h^m)$). Next to accuracy, we also need a numerical method to be efficient with respect to cost, i.e., less consumption of machine time as well as less interference of the user. The other direction is to find out general purpose algorithms which would handle various classes of problems e.g., singular-singularly perturbed problems, singular perturbation problems with turning points and non-linear singular perturbation problems.

Asymptotic expansion techniques, infact, as the literature shows, cover almost all these problem classes. However, for a naive user's point of view, this is a difficult task. For every particular problem, he has to find out the inner expansion and the outer expansion and match them to determine the coefficients of the expansion terms. This matching requires skill, insight into the concerned problem and a great deal of experimentation. An engineer or an applied mathematician needs efficient techniques which he can apply routinely in order to avoid dealing with methods which demand so much of skill, insight and experimentations.

In response to this need for a fresh approach to such problems, some efficient and simple methods are proposed and illustrated in this thesis.

In Chapter 2, a simple decoupling transformation is constructed and analyzed. Unlike the usual Lyapunov type transformation, this transformation partially decouples a singularly perturbed system. The slowly-varying component is decoupled completely, whereas the rapidly varying component involves the slowly varying one. Euler's forward and backward implicit methods are shown to be accurate (in complementary cases) for the rapidly varying component. The non-commutativity of the decouplings and the difference schemes is also shown.

In Chapter 3, an algorithm based on the results of Chapter 2 for linear systems of singularly perturbed two-point boundary-value problems is presented. Numerical experimentations are carried out to demonstrate the effectiveness of the method and thus, the algorithm.

Based on some crucial observations met in the process of demonstrating the effectiveness of the hybrid method, some other numerical methods are presented in Chapter 4. These methods, namely, the Boundary Value Technique and the Cutting Point Technique also take into account the facts known from the asymptotic analysis of problems, where the turning points are absent. These methods also exploit

the numerical boundary layer thickness, an estimate of which is also found out in this chapter. The effectiveness of these special methods are illustrated by examples.

In Chapter 5 the methods are applied to non-linear problems via a modification of Newton's Iteration. A model example of non-linear singular perturbation is taken from Kevorkian and Cole [58] for numerical illustrations.

In Chapter 6, these methods are applied to some optimal control problems, both linear and non-linear.

Finally, in Chapter 7, a veil is drawn as to epitomise this work and scope for further research is also indicated.

In a nut-shell, the numerical methods presented in this thesis for solving systems of singularly perturbed two-point boundary-value problems in ordinary differential equations have been shown to be efficient over the conventional methods. The methods are conceptually simple and need a very modest amount of problem preparation for adapting these to computer implementation. Various kinds of linear and non-linear examples have been solved and numerical results

are presented wherever necessary. It is observed that the accuracy predicted can always be achieved with practically no computational effort. All the numerical results presented in this thesis have been computed on DEC-1090 computer system at IIT Kanpur.

1.6 Preliminary Results

In this section we just introduce some results from various sources, that will be needed in the sequel. The theorems below will appear disconnected from each other; and no attempt is made to connect them. The purpose of adding this section is that it will facilitate presentation of the contents of this thesis. The results given below are non-trivial and they have no obvious connection to Singular Perturbation. Thus, when we refer to these results, it will be convenient for one who does not want to go through the original sources but realise the importance for their applications.

The first theorem tells us about the block-diagonal-ization of a general matrix.

Theorem 6.1 ([38])

Let I be a closed sub-interval of [0,1]. Let the eigenvalues of a time dependent square matrix $A(t)$ of order m be $\lambda_i(t)$, $i = 1,2,\ldots,m$. Let $\varphi(t,\lambda)$ be the polynomial with spectrum $\{ \lambda_i(t) \mid i = 1,2,\ldots,m \}$. Let $\varphi_1(t,\lambda)$, $\varphi_2(t,\lambda)$ be relatively prime polynomials with spectra $\{\lambda_i(t) \mid i = 1,2,\ldots, r\}$ and $\{\lambda_i(t) \mid i = r+1,\ldots,m\}$ respectively. Suppose that $\varphi(t,\lambda) = \varphi_1(t,\lambda) \cdot \varphi_2(t,\lambda)$ for all $t \in I$. Then there is a smooth invertible matrix $M(t)$ of order m, $t \in I$, such that $M^{-1}(t)\, A(t)\, M(t) = \mathrm{diag}(A_1(t), A_2(t))$ where the eigenvalues of $A_1(t)$ and $A_2(t)$ are $\lambda_1(t),\ldots,\lambda_r(t)$, and $\lambda_{r+1}(t),\ldots,\lambda_m(t)$ respectively.

Theorem 6.2 ([68])

Let $\beta_n \geq 0$, $\gamma_n \geq 0$ be constants and consider a grid function v_n satisfying

$$| v_{n+1} | \leq \frac{1}{1+\gamma_n} | v_n | + \frac{\beta_n \gamma_n}{1+\gamma_n} ,$$

for all natural numbers n.

or, for $0 \leq \gamma_n \leq 1$

$$| v_{n+1} | \leq \frac{1-\gamma_n}{1+\gamma_n} |v_n| + \frac{2 \beta_n \gamma_n}{1+\gamma_n} .$$

Then, for positive integer n,

$$| v_n | \leq \max_{0 \leq j \leq n-1} (\beta_j) + \prod_{j=0}^{n-1} \tau_j \, | v_0 |$$

with $\tau_j = \frac{1}{1+\gamma_j}$ for the first case

and $\tau_j = \frac{1 - \gamma_j}{1 + \gamma_j}$ for the second case.

The theorem below is the generalized Taylor's theorem for operators. (Linz [70]).

Theorem 6.3

Let $P : X \longrightarrow Y$ be an operator between two Banach spaces X and Y such that P is n-times continuously differentiable in a neighbourhood $N(x_0,r)$, $r > 0$ of the point x_0 in X. Then, for all x in the interior of $N(x_0,r)$,

$$||P(x) - P(x_0) - \sum_{m=1}^{n-1} \frac{1}{m!} P^{(m)}(x_0) (x-x_0)^m|| \leq \sup_{y \in L(x_0,x)} || P^{(n)}(y) || \frac{||x-x_0||^n}{n!}$$

where $L(x_0,x)$ is the line segment between x_0 and x, and $||.||$ denotes the supremum norm.

Note that the expression $P^{(m)}(x_0)(x-x_0)^m$ is meaningful since $P^{(m)}(x_0)$ is an m-linear operator.

$P^{(3)}(x_o)(x-x_o)^3$, for example, means

$((P^{(3)}(x_o)(x-x_o))(x-x_o))(x-x_o)$.

The next theorem is the Pontryagin's Maximum Principle [95] which can be applied to reduce optimal control problems to a system of differential equations.

Theorem 6.4

Let the scalar cost be defined by $J = \int_{t_o}^{t_1} f^o(x(t),u(t))dt$ which is to be maximized by selecting a control vector $u(t)$ subject to the states $x(t)$ satisfying the differential system $dx/dt = f(x,u)$, $x(t_o) = x_o$. The vector functions $f(x,u)$ and the partial derivatives of each of its components are assumed to be given and are continuous on the direct product $X \times \bar{U}$, where $\bar{U}$ is the closure of U in $\mathbb{R}^r$, the control space and X the state space. Let x^o satisfy the differential equation $\frac{dx^o}{dt} = f^o(x(t),u(t))$ and π be a given line parallel to the x^o axis which passes through the point $(0,x(t_1))$. Let $u(t)$, $t_o \leq t \leq t_1$, be an admissible control such that the corresponding trajectory $x(t)$ which begins at the given point x_o at the time t_o is defined on the entire interval

$t_o \leq t \leq t_1$, through a point on the given line π. In order that $u(t)$ and $x(t)$ be optimal it is necessary that there exists a non-zero absolutely continuous vector function $\Psi(t) = (\Psi_o(t), \ldots, \Psi_n(t))$ corresponding to the functions $u(t)$ and $x(t)$ such that

(i) the function $H(\Psi(t), x(t), u)$

$$= \sum_{j=o}^{n} \Psi_j f^j(x,u)$$ of the variable u in U

attains its maximum at the point $u = u(t)$ almost everywhere in the interval $t_o \leq t \leq t_1$,

(ii) at the terminal time t_1 the relations

$$\Psi_o(t_1) \leq 0, \sup_{u \in U} (\Psi(t_1), x(t_1)) = 0 \qquad (6.1)$$

are satisfied.

Furthermore, it turns out that if $\Psi(t)$, $x(t)$ and $u(t)$ satisfy the system $\dot{x} = H_\Psi$, $\dot{\Psi} = -H_x$ and condition (i) ; then the time function $\Psi_o(t)$ and $\sup_{u \in U} H(\Psi(t), x(t), u)$ are constant. Thus, the above condition (6.1) might be verified at any time t, $t_o \leq t \leq t_1$ and not just at t_1.

The next result, namely, the implicit function theorem is taken from Lang [69].

Theorem 6.5

Let U, V be open sets in Banach spaces E,F respectively, and let $f : U \times V \longrightarrow G$ be a p-times continuously differentiable function, G a Banach space. Let $(a,b) \in U \times V$ and assume that $D_2 f(a,b) : F \longrightarrow G$ is invertibel. (D_2 denotes the partial derivative with respect to the second argument). Suppose that $f(a,b) = 0$. Then there exists a continuous function $g : U_o \longrightarrow V$ defined on an open neighbourhood U_o of the point a such that $g(a) = b$ and $f(x, g(x)) = 0$ for all x in U_o. If U_o is taken to be sufficiently small ball, then g is uniquely determined and also is p-times continuously differentiable.

Next result needs the following definitions [89]. Let $L(R^n,R^m)$ be the vector space of linear operators from R^n to R^m (R^j being j-dimensional real space).

Definition 1

A mapping $F : D \subset R^n \longrightarrow R^m$ is G-differentiable at an interior point x of D if there exists a linear operator $A \in L(R^n,R^m)$ such that for any $h \in R^n$, $\lim_{t \to o} \frac{1}{t} \| F(x+h) - Fx - t\,Ah \| = 0$. The G-derivative of F is denoted by F'.

Definition 2

The mapping $F : D \subset R^n \longrightarrow R^m$ is Frechet differentiable at $x \in \text{Int}(D)$ if there is an $A \in L(R^n, R^m)$ such that

$$\lim_{h \to o} (1/||h||)\ ||F(x+h) - Fx - Ah|| = 0.$$

The linear operator A is called the Frechet-derivative of F at x.

Definition 3

Let $I : x^{k+1} = G\,x^k$, $k = 0,1,\ldots$ be an iterative method (process), where $G : D \subset R^n \longrightarrow R^n$. Then x^* is a point of attraction of the iteration I if there is an open neighbourhood S of x^* such that $S \subset D$ and for any $x^o \in S$, the iterates $\{x^k\}$ defined above all lie in D and converge to x^*.

Definition 4

Let $\{x^k\} \subset R^n$ be any convergent sequence with limit x^*. Then the quantities :

$$Q_p\{x^k\} = \begin{cases} 0, & \text{if } x^k = x^* \text{ for all but finitely many } k. \\ \limsup\limits_{k \to \infty} \dfrac{||x^{k+1} - x^*||}{||x^k - x^*||^p}, & \text{if } x^k \neq x^* \text{ for all but finitely many } k. \\ \infty, & \text{otherwise} \end{cases}$$

Definition 5

Let $C(I,x^*)$ be the set of all sequences with limit x^* generated by an iterative process I. Then,

$$Q_p(I,x^*) = \sup \{ Q_p\, x^* \} \mid \{ x^k \} \in C(I,x^*) \}, \quad 1 \leq p < \infty$$

are the quotient convergence factors of I at x^* with respect to the norm in which the $Q_p\{ x^k \}$ are computed.

$$O_Q(I,x^*) = \begin{cases} \infty, & \text{if } Q_p(I,x^*) = 0, \quad \forall p \in [1,\infty) \\ \inf \{ p \in [1,\infty) \mid Q_p(I,x^*) = \infty \}, & \text{otherwise} \end{cases}$$

is the Q-order of I at x^*.

Definition 6

Let $\{ x^k \} \subset R^n$ be any sequence that converges to x^*. Then the numbers

$$R_p\{ x^k \} = \begin{cases} \limsup\limits_{k\to\infty} ||x^k - x^*||^{1/k}, & \text{if } p = 1 \\ \limsup\limits_{k\to\infty} ||x^k - x^*||^{1/k^p}, & \text{if } p > 1 \end{cases}$$

are the root convergence factors of the sequence $\{x^k\}$. If I is an iterative process with limit point x^* and $C(I,x^*)$ is the set of all sequences generated by I which converge to x^*, then

$$R_p(I,x^*) = \sup \{R_p\{x^k\} \mid \{x^k\} \in C(I,x^*) , \quad 1 \leq p < \infty$$

are the R-factors of I at x^*.

The quantity

$$O_R(I, x^*) = \begin{cases} \infty, & \text{if } R_p(I,x^*) = 0, \ \forall p \in [1,\infty) \\ \inf\{ p \in [1,\infty) \mid R_p(I,x^*) = 1\}, & \text{otherwise;} \end{cases}$$

is called the R-order of I at x^*.

Now we state the convergence theorem of the Newton's iteration [89, Thm 10.22].

Theorem 6.6

Assume that $F : D \subset R^n \longrightarrow R^n$ is G-differentiable on an open neighbourhood $S_o \subset D$ of a point $x^* \in D$ for which $Fx^* = 0$, and that F' is continuous at x^* and $F'(x^*)$ non-singular. Then x^* is a point of attraction of the Newton iteration

$I : x^{k+1} = x^k - F'(x^k)^{-1} F x^k$, $k = 0,1,\ldots$ and $R_1(I,x^*) = Q_1(I,x^*) = 0$. If, in addition, there are constants $\alpha < \infty$ and $p \in (0,1]$ such that $||F'(x) - F'(x^*)|| \leq \alpha |x-x^*|^p$, $\forall x \in S_o$, then $O_R(I,x^*) \geq O_Q(I,x^*) \geq 1+p$. Finally, if F is continuously Frechet-differentiable on S_o and the second Frechet-derivative of F exists at x^* and satisfies $F''(x^*)h \neq 0$, $\forall h \in R^n$, $h \neq 0$ then $O_R(I,x^*) = O_Q(I,x^*) = 2$; where the last F'' denotes the second Frechet derivative of F.

Chapter 2

A METHOD FOR LINEAR SINGULARLY PERTURBED SYSTEMS

2.1 Introduction

For numerical integration of ordinary differential equations, the first and foremost task is to find out an appropriate grid. This appropriateness can be best represented via the assumptions on the coefficient matrix and the non-homogeneous term, if any, occuring in the differential equation. This is exactly done, for example, in Kreiss et al. [68] who prove the boundedness of the solution and its derivatives upon assuming some numerically verifiable conditions on the system matrix and the non-homogeneous term for stiff linear systems. Our concern here is a special subclass of these stiff systems, namely, the singularly perturbed systems. Ferguson [32] considers the two-point boundary-value systems of the type

$$\begin{bmatrix} \varepsilon\,\dot{y}_1 \\ \dot{y}_2 \\ \varepsilon\,\dot{y}_3 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \varepsilon A_{13} \\ A_{21} & A_{22} & A_{23} \\ \varepsilon A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}, 0 < t < 1 \qquad (1.1a)$$

with the boundary conditions

$$B_0\, y(0) + B_1\, y(1) = B_2 \qquad (1.1b)$$

where $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, (1.1b) represents n-linearly independent boundary conditions, $y \in \mathbb{R}^n$; $0 < \varepsilon << 1$, a parameter, '.' denotes d/dt. Assuming that the matrices $A_{ij}(t)$, $f_i(t)$ are smooth (possess continuous derivatives upto some finite order) functions of t satisfying the stability conditions

$$\text{Re}\,\lambda(A_{11}) < -\mu < 0 < \mu < \text{Re}\,\lambda(A_{33})$$

he proves the uniqueness of a solution y which admits of the following a priori bounds :

$$||y_1|| \leq K(|y_1(0)| + |y_2(0)| + \varepsilon|y_3(1)| + \sum_{i=1}^{3} ||f_i||) \tag{1.2a}$$

$$||y_2|| \leq K(\varepsilon|y_1(0)| + |y_2(0)| + \varepsilon|y_3(1)| + \sum_{i=1}^{3} ||f_i||) \tag{1.2b}$$

$$||y_3|| \leq K(\varepsilon|y_1(0)| + |y_2(0)| + \varepsilon|y_3(1)| + \sum_{i=1}^{3} ||f_i|| \tag{1.2c}$$

with $||g|| = \int_0^1 |g(t)|dt$, and $K(> 0)$ is a real constant. The results corresponding to (1.2) for the general class of linear singularly perturbed two-point boundary value systems :

$$\begin{bmatrix} \dot{y}_1 \\ \varepsilon\dot{y}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad 0 < t < 1 \tag{1.3}$$

with similar boundary conditions (1.1b) are also given in his thesis [31].

In addition to these a priori estimates of the solution, we also know the geometric behaviour of the solution from the asymptotic analysis. In general, the boundedness results are proved under the conditional stability criterion, that the eigenvalues of the system matrix $A(t,\varepsilon) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ are bounded away from the imaginary axis. Under this conditional stability criterian, the rapidly varying component $y_2(t,\varepsilon)$ of the solution $y(t,\varepsilon)$ can become a combination of two pieces contained in different manifolds. The dichotomies (ordinary or exponential) present in the fast varying component gives the motivation to decouple the system first into two modes: slowly varying and the rapidly varying one. From Coppel [20-21] the theorem below follows ; (Kokotovic et al. [62, Ch-5]).

Theorem 1.1

Let $B(t)$ be a continuously differentiable matrix function such that for every $t \in [0,1]$, $||B(t)|| \leq M$, $||\dot{B}(t)|| \leq \delta$, and

there exists $\alpha > 0$: Re $\lambda(B)(t)) \leq -\alpha$. (1.4)

Then there exists $\varepsilon_o > 0$ such that, for $0 < \varepsilon < \varepsilon_o$ the system $\varepsilon \dot{z} = B(t)z$ has a fundamental matrix $X(t)$ satisfying the inequality

$$||X(t)\, X^{-1}(s)|| \leq K\, e^{-\alpha(t-s)} \tag{1.5}$$

for $t \geq s$, for some const. $K > 0$. Moreover, if condition (1.4) is replaced by

there exists $\alpha > 0$: Re $\lambda(B) \geq \alpha$, (1.6)

then the inequality

$$||X(t)X^{-1}(s)|| \leq K\, e^{-\alpha(s-t)} \tag{1.7}$$

for $s \geq t$, holds for some constant $K > 0$.

Remark 1.1

The constants ε_o and K, in case (1.6) is satisfied (and thus in (1.7)) might differ from those for the case where (1.4) is satisfied (and accordingly in (1.5)). For the second case (i.e., (1.6)) we write the corresponding ε_o and K by ε_1 and K_o respectively.

In the following we abbreviate the above matrix product $X(t)X^{-1}(s)$ by refering it to as 'the matrizant'

$\varphi(t,s) = X(t)\, X^{-1}(s)$; and

$\varphi(s,t) = X(s)\, X^{-1}(t)$. As such the inequalities (1.5) and (1.7) will be interpreted as giving estimates of the matrizants.

2.2 A Decoupling Transformation

The usual way of decoupling the system

$$\begin{bmatrix} \dot{y}_1 \\ \varepsilon \dot{y}_2 \end{bmatrix} = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} \quad (2.1)$$

for $0 < t < 1$, is done by finding a decoupling Lyapunov transformation $\tilde{T}(t) = \begin{bmatrix} I_m & \varepsilon \tilde{H}(t) \\ -\tilde{L}(t) & I_k - L(t)\tilde{H}(t) \end{bmatrix}$ ($y_1 \in \mathbb{R}^m$, $y_2 \in \mathbb{R}^k$) which puts (2.1) in the form

$$\dot{\xi}(t) = (A_{11}(t) - A_{12}(t)\, \tilde{L}(t))\, \xi(t) + g_1(t) \quad (2.2a)$$

$$\varepsilon \dot{\eta}(t) = (A_{22}(t) + \varepsilon \tilde{L}(t)\, A_{12}(t))\, \eta(t) + g_2(t) \quad (2.2b)$$

with $\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \tilde{T}^{-1}(t) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$.

But this involves in solving two Riccati differential equations given by

$$\varepsilon \dot{\tilde{L}} = A_{22}\tilde{L} - A_{21} - \varepsilon \tilde{L}(A_{11} - A_{12})\tilde{L} \quad (2.3a)$$

$$-\varepsilon \dot{\tilde{H}} = \tilde{H}A_{22} - A_{12} + \varepsilon \tilde{H}LA_{12} - \varepsilon(A_{11} - A_{12}\tilde{L})\tilde{H}\,. \quad (2.3b)$$

Again, since the slow state y_1 (or a linear combination of y_1 and y_2) gives no trouble (analytically or numerically), we need (for our purpose) only separate the slow state and allow the differential equation for the fast state y_2 to contain the slow state variables (see Matheij [74]). To this end, we assume the following :

(A-1). The matrices $A_{ij}(t)$, $i,j = 1,2$ are continuously differentiable.

(A-2). There are constants $c_{ij}^1 > 0$ and $c_{ij}^2 > 0$ such that $||A_{ij}(t)|| \leq c_{ij}^1$, $||\dot{A}_{ij}(t)|| \leq c_{ij}^2$

(A-3). There are constants $c_3 > 0$ and $c_4 > 0$ such that any eigenvalue $\lambda(t)$ of $A_{22}(t)$ satisfies :

Re $\lambda(t) \leq - c_3$ or Re $\lambda(t) \geq c_4$.

Under these assumptions we will establish the existence and boundedness of a transformation

$$T = \begin{bmatrix} I_m & \varepsilon L \\ O & I_k \end{bmatrix} \qquad (2.4)$$

which decouples the slow variable from the system (2.1).

Let $E = \begin{bmatrix} I_m & O \\ O & \varepsilon I_k \end{bmatrix}$,

then $E^{-1} = \begin{bmatrix} I_m & 0 \\ 0 & \varepsilon^{-1} I_k \end{bmatrix}$.

Assuming that L is bounded T^{-1} admits of the following representation.

$$T^{-1} = \begin{bmatrix} I_m & -\varepsilon L \\ 0 & I_k \end{bmatrix} .$$

Denote $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$.

The system (2.1) can be rewritten as

$$E\dot{y} = Ay + f \tag{2.5}$$

$$\Longleftrightarrow \quad \dot{y} = E^{-1}Ay + E^{-1}f$$

$$\Longleftrightarrow \quad ET^{-1}\dot{y} = ET^{-1}E^{-1}ATT^{-1}y + ET^{-1}E^{-1}f$$

$$\Longleftrightarrow \quad ET^{-1}\dot{y} - ET^{-1}\dot{T}T^{-1}y = ET^{-1}(E^{-1}AT - \dot{T})T^{-1}y + ET^{-1}E^{-1}f$$

$$\Longleftrightarrow \quad E\frac{d}{dt}(T^{-1}y) = ET^{-1}(E^{-1}AT - \dot{T})(T^{-1}y) + (ET^{-1}E^{-1}f) \tag{2.6}$$

Substitute $z = T^{-1}y$ (2.7a)

$$g = ET^{-1}E^{-1}f \tag{2.7b}$$

in (2.6) to obtain the equivalent system to (2.5) and hence to (2.1) :

$$\dot{Ez} = \bar{A}z + g \,, \quad \bar{A} = ET^{-1}(\dot{E}^{-1}AT - \dot{T}) \tag{2.8}$$

Now,

$$ET^{-1}\,E^{-1} = \begin{bmatrix} I_m & -L \\ O & I_k \end{bmatrix} \tag{2.9a}$$

and

$$\bar{A} = \begin{bmatrix} A_{11}-LA_{21} & \varepsilon A_{11}L+A_{12}-\varepsilon\dot{L}-\varepsilon LA_{21}L-LA_{22} \\ A_{21} & \varepsilon A_{21}L + A_{22} \end{bmatrix} \tag{2.9b}$$

In the system (2.8), the slow variable is decoupled from the system, if and only if the upper-right block in $\bar{A}$ is zero; i.e., we put

$$\varepsilon\dot{L} = -L\,A_{22} + A_{12} + \varepsilon(A_{11} - L\,A_{21})L \tag{2.10}$$

Our task now is to show that for sufficiently small ε, there exists a bounded continously differentiable matrix $L(t)$ satisfying (2.10). We choose the boundary conditions for (2.10) in such a way that the boundary layers in its solution are eliminated. Thus we are justified in seeking a solution of (2.10) in (an asymptotic) series form. However, the eigenvalues of $A_{22}(t)$, the system matrix of (2.10), are not grouped, i.e., with respect to their signs of real parts; which is necessary for the dichotomies involved. Thus before seeking an asymptotic solution to (2.10), we first order (regroup) the eigenvalues.

From (A-3) and the boundedness of A_{22}, $\dot{A}_{22}$ it follows that (see Coppel [20], Gingold [38]) there exists a non-singular continuously differentiable bounded matrix $W(t)$ with bounded derivative such that

$$W(t)\, A_{22}(t)\, W^{-1}(t) = D(t) = \begin{bmatrix} D_1(t) & 0 \\ 0 & D_2(t) \end{bmatrix} \tag{2.11a}$$

where

$$\mathrm{Re}\,\lambda\,(D_1(t)) \leq -c_3 < 0 \tag{2.11b}$$

$$\mathrm{Re}\,\lambda\,(D_2(t)) \geq c_4 > 0 \,. \tag{2.11c}$$

Regarding this, we let

$$L = MW \tag{2.12}$$

to reduce (2.10) to :

$$\varepsilon \dot{M} = -MD + A_{12}W^{-1} + \varepsilon(A_{11} - MWA_{21})M - \varepsilon \dot{M}\dot{W}W^{-1} \tag{2.13}$$

We seek a solution of (2.13) in the form :

$$M(t) = \sum_{j=0}^{N-1} \varepsilon^j M_j(t) + \varepsilon^N R(t) \tag{2.14}$$

In fact we are interested in an approximation $O(\varepsilon^2)$, i.e., $N = 2$.

Substituting (2.14) in (2.13) and comparing like powers of ε, we obtain :

$$M_o = A_{12}W^{-1}D^{-1} = A_{12}A_{22}^{-1}W^{-1} \qquad (2.15a)$$

$$M_1 = (-\dot{M}_o + A_{11}M_o - M_oWA_{21}M_o - M_o\dot{W}W^{-1})D^{-1} \qquad (2.15b)$$

and so on.

In general, the remainder term R satisfies an equation of the form :

$$\varepsilon\dot{R} = -RD - \dot{M}_{N-1} + f_1(M_o,\ldots,M_{N-1}) + \varepsilon f_2(R, M_o,\ldots, M_{N-1}, \varepsilon). \qquad (2.16)$$

Using (2.16), we can get an approximation of any order, i.e., $O(\varepsilon^N)$ for any N. We are interested only in N = 2. In this case,

$$f_1 \equiv (A_{11} - M_oWA_{21})M_1 - M_1(WA_{21}M_o + \dot{W}W^{-1}) \qquad (2.17a)$$

$$f_2(R) \equiv (A_{11} - M_oWA_{21} - \varepsilon M_1WA_{21})R - RWA_{21}(M_o + \varepsilon M_1 + \varepsilon^2R) - \varepsilon R\dot{W}W^{-1} \qquad (2.17b)$$

and (2.16) reduces to

$$\varepsilon\dot{R} = -RD - \dot{M}_1 + f_1 + \varepsilon f_2 \qquad (2.18)$$

The above discussion shows that our aim is accomplished provided that the solution R of (2.18) remains bounded.

Let $\varphi(t,s)$ be the matrizant of the system $\dot{x} = D_1(t)x$ and $\Psi(t,s)$ be the matrizant of the system $\dot{x} = D_2(t)x$.

Let $N_o(t) = -\dot{M}_1(t) + f_1$ (2.19a)

$N(t) = N_o(t) + \varepsilon f_2$. (2.19b)

Define an integral operator $S : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$S(t) = \frac{1}{\varepsilon}\int_o^t N(s)\begin{bmatrix}\varphi(s,t) & 0\\ 0 & 0\end{bmatrix} ds - \frac{1}{\varepsilon}\int_t^1 N(s)\begin{bmatrix}0 & 0\\ 0 & \Psi(s,t)\end{bmatrix} ds \qquad (2.20)$$

Lemma 2.1

Let $\bar{R} = \{R : ||R|| < \rho\}$ for some $\rho > 0$. If (A-1) - (A-3) are satisfied then, there are constants $K_1, K_2, K_3, K_4, \varepsilon_1$ such that

(a) $|| N_o(t) || \leq K_1$,

(b) $|| N(t) || \leq K_1 + \varepsilon\rho K_2 + \varepsilon^2\rho^2 K_3$,

(c) $|| f_2(R_1) - f_2(R_2) || \leq K_4 || R_1 - R_2 ||$ for $R_1, R_2 \in \bar{R}$ and $0 < \varepsilon < \varepsilon_1$,

(d) $R \in \bar{R}$ is a solution of (2.18) iff it satisfies the integral equation $R = S$.

Proof

(a) - (b) follow from the boundedness of $A_{ij}(t)$ and $\dot{A}_{ij}(t)$, $M_o(t)$, $M_1(t)$ on $[0,1]$. (c)

follows from using the matrix identity

$$R_1WA_{21}R_1 - R_2WA_{21}R_2$$
$$= (R_1 - R_2)WA_{21}R_1 + R_2WA_{21}(R_1 - R_2)$$

since, $f_2(R_1) - f_2(R_2)$

$$= (A_{11} - M_oWA_{21} - \varepsilon M_1WA_{21})R_1$$
$$- R_1WA_{21}(M_o + \varepsilon M_1 + \varepsilon^2 R_1) - \varepsilon R_1\dot{W}W^{-1}$$
$$- [(A_{11} - M_oWA_{21} - \varepsilon M_1WA_{21})R_2$$
$$- R_2WA_{21}(M_o + \varepsilon M_1 + \varepsilon^2 R_2) - \varepsilon R_2\dot{W}W^{-1}]$$

$$= (A_{11} - M_oWA_{21} - \varepsilon M_1WA_{21})(R_1 - R_2)$$
$$- (R_1 - R_2)[WA_{21}(M_o + \varepsilon M_1) + \varepsilon\dot{W}W^{-1}]$$
$$- \varepsilon^2[R_1WA_{21}R_1 - R_2WA_{21}R_2] ,$$

and $\dot{W}$ as well as W^{-1} are bounded. The 'if' part in (d) follows from differentiation and the 'only if' part that from variation of constants formula.

Lemma 2.2

If (A-1) - (A-3) are satisfied then there exists $\varepsilon^* > 0$ such that the integral equation $R = S$ has a unique solution in $\bar{R}$, for $0 < \varepsilon < \varepsilon^*$.

Proof

By contraction mapping principle, the lemma

will be proved if S happens to be a contraction mapping.

Now, $S(R_1) - S(R_2)$

$$= \frac{1}{\varepsilon} \int_0^1 [N_0 + \varepsilon f_2(R_1)] \begin{bmatrix} \varphi(s,t), & 0 \\ 0 & 0 \end{bmatrix} ds$$

$$- \frac{1}{\varepsilon} \int_t^1 [N_0 + \varepsilon f_2(R_1)] \begin{bmatrix} 0 & 0 \\ 0 & \Psi(s,t) \end{bmatrix} ds$$

$$- \frac{1}{\varepsilon} \int_0^1 [N_0 + \varepsilon f_2(R_2)] \begin{bmatrix} \varphi(s,t) & 0 \\ 0 & 0 \end{bmatrix} ds$$

$$+ \frac{1}{\varepsilon} \int_t^1 [N_0 + \varepsilon f_2(R_2)] \begin{bmatrix} 0 & 0 \\ 0 & \Psi(s,t) \end{bmatrix} ds$$

$$= \frac{1}{\varepsilon} \int_0^1 \varepsilon [f_2(R_1) - f_2(R_2)] \begin{bmatrix} \varphi(s,t) & 0 \\ 0 & 0 \end{bmatrix}$$

$$- \frac{1}{\varepsilon} \int_t^1 \varepsilon [f_2(R_1) - f_2(R_2)] \begin{bmatrix} 0 & 0 \\ 0 & \Psi(s,t) \end{bmatrix} ds$$

since N_0 does not depend upon R. Hence from lemma 2.1 and theorem 1.1 it follows that there exists $\varepsilon_2 > 0$ such that, for $0 < \varepsilon < \varepsilon_2 \leq \min(\varepsilon_0, \varepsilon_1)$

$$||S(R_1) - S(R_2)||$$

$$\leq \left[\int_0^t K\, e^{-c_3(t-s)/\varepsilon}\, ds + \int_1^t K_o\, e^{-c_4(t-s)/\varepsilon} ds \right] \times K_4\, ||\, R_1 - R_2\, ||$$

$$\leq \frac{\varepsilon \max(K,K_o).K_4}{\min(c_3,c_4)} \cdot ||R_1 - R_2|| \,.$$

Hence,

$$||\, S(R_1) - S(R_2)\, ||\ \leq \tfrac{1}{2}\, ||\, R_1 - R_2\, ||$$

$$\text{for}\quad \varepsilon < \varepsilon_3 = \frac{\min\,(c_3,\, c_4)}{2\, K_4\, \max(K,K_o)} \,. \qquad (2.21)$$

Also, from Lemma 2.1, $||\, S\, ||$

$$\leq \frac{1}{\varepsilon} \left[\int_0^t K\, e^{-c_3(t-s)/\varepsilon} ds + \int_1^t K_o\, e^{-c_4(t-s)/\varepsilon} ds \right] \times (K_1 + \varepsilon\rho K_2 + \varepsilon^2\rho^2 K_3)$$

$$\leq \frac{\max(K,K_o)}{\min(c_3,c_4)} (K_1 + \varepsilon\rho K_2 + \varepsilon^2\rho^2 K_3). \qquad (2.22)$$

We choose

$$\rho = \frac{2\max(K,K_o).K_1}{\min(c_3,\, c_4)} \,,$$

$$\varepsilon^* = \min\left(\varepsilon_o,\ \varepsilon_1,\ \frac{\min(c_3,c_4)}{3K_2\ \max(K,K_o)},\ \frac{K_2}{2K_4\rho},\ \frac{\min(c_3,\ c_4)}{2K_4\ \max(K,K_o)}\right).$$

Then, from (2.21) and (2.22) it follows that $||S|| < \rho$ and S is a contraction mapping on $\bar{R}$.

From the discussions above and the lemmas 2.1 and 2.2, we obtain the following theorem :

Theorem 2.1

Assume that the system (2.1) satisfies the assumptions (A-1) - (A-3). Then there exist an $\varepsilon^* > 0$ and continuously differentiable matrices $W(t), M_o(t), M_1(t)$ bounded on $[0,1]$ such that

$$W(t)\ A_{22}(t)W^{-1}(t) = \mathrm{diag}(D_1(t),\ D_2(t))$$

$$\mathrm{Re}\,\lambda(D_1) \le -c_3\ , \qquad \mathrm{Re}\,\lambda(D_2) \ge c_4$$

and for $0 < \varepsilon < \varepsilon^*$,

$$M(t) = M_o(t) + \varepsilon M_1(t) + O(\varepsilon^2)$$

the system (2.1) is transformed to

$$\dot{z}_1 = [A_{11}(t) - M(t)W(t)A_{21}(t)]z_1 + g_1(t)$$

$$\varepsilon\dot{z}_2 = A_{21}(t)z_1 + [\varepsilon A_{21}(t)M(t)W(t) + A_{22}(t)]z_2 + g_2(t)$$

$$y_1 = z_1 + M(t)W(t)z_2$$

$$y_2 = z_2$$

Lemma 3.1

Consider the system

$$\varepsilon\dot{x}(t) = [B(t) + \varepsilon^{\gamma} \Gamma(t,\varepsilon)]x(t) \qquad (3.1)$$

where $||\Gamma(t,\varepsilon)|| \leq c_5$, for every $t \in [0,1]$ and $\gamma > 0$. Let $\varphi_1(t,s)$ be the matrizant of (3.1). Corresponding to this consider also the system

$$\varepsilon\dot{x}(t) = B(t)x(t) \qquad 0 < t < 1 \qquad (3.2)$$

Let $\varphi(t,s)$ be the matrizant of (3.2). Assume that $||B(t)||$, $||\dot{B}(t)||$ are bounded and $\mathrm{Re}\,\lambda(B(t)) < -\alpha < 0$. Then there exist positive constants ε^*, α_1 and K_5 such that for all ε, $0 < \varepsilon < \varepsilon^*$,

$$||\varphi_1(t,s) - \varphi(t,s)|| \leq \varepsilon^{\gamma} K_5 e^{-\alpha_1(t-s)/\varepsilon}, \quad \text{for } t \geq s. \qquad (3.3)$$

Due to the exponential dichotomy (Coppell [21]) if the stability condition is replaced by

$$\mathrm{Re}\,\lambda(B(t)) > \alpha_2 > 0,$$

then, the inequality corresponding to (3.3) will be

$$||\varphi_1(t,s) - \varphi(t,s)|| \leq \varepsilon^{\gamma} K_5 e^{-\alpha_2(s-t)/\varepsilon} \quad \text{for } t \leq s. \qquad (3.4)$$

Due to Lemma 3.1, it suffices to consider the fast

system of the form

$$\varepsilon \dot{x} = B(t)x + f(t) \qquad 0 < t < 1 \tag{3.5}$$

Assume that $B(t)$ is conditionally stable and $||B(t)||$, $||B^{-1}(t)||$, $||f(t)||$ are bounded. Let $\{h_n\}_{n-1}$ be a grid defined on $[0,1]$, $h = \max_n\{h_n\}$, $t_n = \sum_{j=o} h_j$, $t_o = 0$, $t_N = 1$, $n = 1,\ldots,N-1$. Assume that

$$\min_n\{h_n\} > \max\left\{\frac{\varepsilon}{|B|}, 2\varepsilon|B|\right\}. \tag{3.6}$$

In fact, for standard methods to be accurate for the original system (2.1), one has to take, in general $h << \varepsilon$, which is impracticable. Again this type of grid-generation with $h<<\varepsilon$ also affects the solution causing more round-off errors. Hence we are justified in assuming the minimum of the grid lengths to be greater than some number of $O(\varepsilon)$. However (3.6) permits us to take very small grid lengths, since $\varepsilon << 1$. Let the numerical solution on the grids be denoted by $u_n = u(t_n)$ and similarly $B(t_n) = B_n$. In fact, in the course of error estimation we only need an assumption weaker than (3.6), namely,

$$2\frac{\varepsilon}{h_n} < |B_{n+1}^{-1}| < \frac{h_n}{\varepsilon} \tag{3.6'}$$

We will apply Euler's Implicit Scheme to solve systems of the form (3.5). This particular method is chosen due to its unconditional contractivity proerty. This property holds when the one-sided Lipschitz constant $\nu(t) \leq 0$ (see Dekker et al. [22] , see also Kreiss et al. [68]).

The difference equation corresponding to (3.5) is :

$$\frac{\varepsilon}{h_n} (u_{n+1} - u_n) = B_{n+1} u_{n+1} + f_{n+1}$$

$$\text{or,} \quad -\frac{\varepsilon}{h_n} u_n = (B_{n+1} - \frac{\varepsilon I}{h_n}) u_{n+1} + f_{n+1} . \qquad (3.7)$$

Due to the conditional stability, B_{n+1} is non-singular; thus (3.7) can be rewritten by

$$-\frac{\varepsilon}{h_n} u_n = B_{n+1}(I - \frac{\varepsilon}{h_n} B_{n+1}^{-1}) u_{n+1} + f_{n+1}$$

In view of (3.6) or (3.6'), the matrix

$I - \frac{\varepsilon}{h_n} B_{n+1}^{-1}$ is invertible. Hence the difference scheme (3.7) is equivalent to :

$$u_{n+1} = -\frac{\varepsilon}{h_n} (I - \frac{\varepsilon}{h_n} B_{n+1}^{-1})^{-1} B_{n+1}^{-1} u_n$$

$$- (I - \frac{\varepsilon}{h_n} B_{n+1}^{-1})^{-1} B_{n+1}^{-1} f_{n+1} \qquad (3.8)$$

But $|(I - \frac{\varepsilon}{h_n} B_{n+1}^{-1})^{-1}|$

$= |I + \frac{\varepsilon}{h_n} B_{n+1}^{-1} + (\frac{\varepsilon}{h_n} B_{n+1}^{-1})^2 + \ldots|$ (see [54])

$\leq 1 + |\frac{\varepsilon}{h_n} B_{n+1}^{-1}| + |\frac{\varepsilon}{h_n} B_{n+1}^{-1}|^2 + \ldots$

$= (1 - |\frac{\varepsilon}{h_n} B_{n+1}^{-1}|)^{-1}$ (due to (3.6')) .

Hence (3.8) implies that

$$|u_{n+1}| \leq \frac{\varepsilon}{h_n} \frac{|B_{n+1}^{-1}|}{1 - \frac{\varepsilon}{h_n}|B_{n+1}^{-1}|} |u_n| +$$

$$+ \frac{|B_{n+1}^{-1}|}{1 - \frac{\varepsilon}{h_n}|B_{n+1}^{-1}|} |f_{n+1}|$$

Letting $\alpha_{n+1} = |B_{n+1}^{-1}|^{-1} = O(1)$ (3.9a)

$\theta = \varepsilon/h_n < 1 \quad (<< 1)$. (3.9b)

the above inequality reduces to

$$|u_{n+1}| \leq \frac{\theta}{\alpha_{n+1} - \theta}|u_n| + \frac{|f_{n+1}|}{\alpha_{n+1} - \theta} \quad . \tag{3.9c}$$

Let $\gamma_n = \frac{\alpha_{n+1}}{\theta} - 2 \geq 0$ (3.10a)

(due to 6' or 6)

$0 \leq \beta_n = \frac{|f_{n+1}|}{\alpha_{n+1} - 2\theta} = O(1)$ (3.10b)

Then $$\frac{1}{1+\gamma_n} = \frac{1}{1 + \frac{\alpha_{n+1}}{\theta} - 2} = \frac{\theta}{\alpha_{n+1} - \theta}$$

$$\frac{\gamma_n \beta_n}{1+\gamma_n} = \frac{(\frac{\alpha_{n+1}}{\theta} - 2)\frac{|f_{n+1}|}{\alpha_{n+1} - 2\theta}}{1 + \frac{\alpha_{n+1}}{\theta} - 2}$$

$$= \frac{|f_{n+1}|/\theta}{(\alpha_{n+1} - \theta)/\theta} = \frac{|f_{n+1}|}{\alpha_{n+1} - \theta}$$

Hence (3.9c) can be written as :

$$|u_{n+1}| \leq \frac{1}{1+\gamma_n} |u_n| + \frac{\gamma_n \beta_n}{1+\gamma_n} \quad . \tag{3.11}$$

Applying theorem 6.2 in Chapter 1 we obtain

$$|u_n| \leq \max_{0 \leq j \leq n-1} \beta_j + \left(\prod_{j=0}^{n-1} \frac{1}{1+\gamma_j}\right) |u_o| \quad . \tag{3.12}$$

But $$\beta_j = \frac{|f_{j+1}|}{|B_{j+1}^{-1}|^{-1} - 2\varepsilon/h_j}$$

$$\frac{1}{1 + \gamma_j} = \frac{\varepsilon/h_j}{|B_{j+1}^{-1}|^{-1} - \varepsilon/h_j}$$

Hence,

$$|u_n| \leq \max_{0 \leq j \leq n-1} \frac{|f_{j+1}|}{|B_{j+1}^{-1}|^{-1} - 2\varepsilon/h_j}$$

$$+ \left(\prod_{j=0}^{n-1} \frac{\varepsilon/h_j}{|B_{j+1}^{-1}|^{-1} - \varepsilon/h_j}\right) |u_o| \tag{3.13}$$

which shows that $|u_n|$ is bounded. Similarly, the error-estimate can also be derived. Let the solution $x(t)$ be twice continuously differentiable, or, in other words, $B(t)$ and $f(t)$ are so. Let us denote $x_n = x(t_n)$, the exact solution at the grid points t_n's. By Taylor's theorem,

$$\frac{\varepsilon}{h_n}(x_{n+1} - x_n) = B_{n+1} x_{n+1} + f_{n+1} + \tau_{n+1}$$

with $|\tau_{n+1}| \leq \frac{1}{2} h_n \max_{t_n \leq \eta \leq t_{n+1}} |\ddot{x}(\eta)|$.

Then $e_n = x_n - u_n$ satisfies :

$$\frac{\varepsilon}{h_n}(e_{n+1} - e_n) = B_{n+1} e_{n+1} + \tau_{n+1} \qquad (3.14)$$

Now, (3.14) is of the same form as (3.7); analogously it follows that

$$|e_n| \leq \max_{0 \leq j \leq n-1} \frac{\tau_{j+1}}{|B_{j+1}^{-1}|^{-1} - 2\varepsilon/h_j} + \prod_{j=0}^{n-1} \frac{\varepsilon/h_j}{|B_{j+1}^{-1}|^{-1} - \varepsilon/h_j} |e_o|$$

$$\leq \max_{0 \leq j \leq n-1} \frac{h_j \tau_{j+1}}{h_j |B_{j+1}^{-1}|^{-1} - 2\varepsilon} + \prod_{j=0}^{n-1} \frac{\varepsilon |e_o|}{h_j |B_{j+1}^{-1}|^{-1} - \varepsilon} . \qquad (3.15)$$

Writing $e = \max_n |e_n|$, we obtain,

$$e = \begin{cases} O(h^2), & \text{if } h||B|| >> 1 \\ O(h), & \text{otherwise} \end{cases}$$

We summarise the above discussion in the following theorem.

Theorem 3.1

Let $B(t)$, $f(t)$ be twice continuously differentiable; $||B(t)||$, $||f(t)||$ are bounded. Let h_n be a grid satisfying (3.6) (or (3.6')). Then the difference scheme (3.7) solves (3.5) with u_n bounded and error

$e = O(h^2)$ if $h||B|| >> 1$; otherwise $e = O(h)$.

Remark 3.1

If instead of Euler's Implicit scheme, we apply Euler's Explicit scheme and integrate backward, i.e., use the scheme

$$(B_n + \frac{\varepsilon}{h_n} I)u_n = \frac{\varepsilon}{h_n} u_{n+1} - f_n \ , \qquad (3.16)$$

then as previously, we get the estimate (analogous to (3.13))

$$|u_{N-n}| \leq \max_{0\leq j\leq n-1} \frac{|f_{N-j}|}{|B_{N-j}^{-1}|^{-1} + 2\varepsilon/h_{N-j}} \cdot$$

$$+ \prod_{j=0}^{n-1} \frac{\varepsilon/h_{N-j}}{|B_{N-j}^{-1}|^{-1} + \varepsilon/h_{N-j}} |u_N| \qquad (3.17)$$

and the corresponding error estimate (analogous to (3.15)) :

$$|e_{N-n}| \leq \max_{0\leq j\leq n-1} \frac{h_{N-j}\ \tau_{N-j-1}}{h_{N-j}|B_{N-j}^{-1}|^{-1} + 2\varepsilon} +$$

$$+ \prod_{j=0}^{n-1} \frac{|e_N|}{h_{N-j}|B_{N-j}^{-1}|^{-1} + \varepsilon} \cdot \qquad (3.18)$$

Again, here $e = O(h^2)$ if $h||B|| >> 1$,
otherwise $e = O(h)$.

Remark 3.2

A look at Theorem 1.1 reveals the fact that if $\mathrm{Re}\,\lambda(B(t)) < 0$, the forward integration is stable, whereas, if $\mathrm{Re}\,\lambda(B(t)) > 0$, then the stability holds in the negative direction; suggesting backward integration. Due to this reason, any numerical method which tries the solution of (3.1) should

integrate forward or backward depending upon the sign of the real parts of the eigenvalues of the system matrix $B(t)$.

The above error estimates and Remark 3.2 suggest that we employ the Euler's implicit method and integrate forward when the eigenvalues of $B(t)$ have negative real parts and employ Euler's Explicit method backward when the eigenvalues of $B(t)$ have positive real parts. If $B(t)$ have eigenvalues some with positive and some with negative real parts, a transformation (as in equation (2.11)) of $B(t)$ must be done to separate the eigenvalues with negative or positive real parts before applying the methods. Note that for such a transformation it is necessary to assume that on the whole of the interval $[0,1]$, the spectrum can be divided, i.e., any eigenvalue $\bar{\lambda}(t)$ of $B(t)$ must be either having real part negative or positive throughout the interval. In practice, very often the contrary happens, i.e., $\mathrm{Re}\,(\bar{\lambda}(t))$ changes sign, when t varies through $[0,1]$. To overcome this, we assume a weaker statement to hold. Let $\{b_i\}$ be a partition of $[0,1]$:

$0 = b_0 < b_1 < \ldots < b_k = 1.$

Denote $\beta_i = [b_i, b_{i+1})$, $i = 1,\ldots,k-2$; $\beta_{k-1} = [b_{k-1},1]$.

(A-4) Assume that on each sub-interval β_i the eigenvalue criterion, i.e., any eigenvalue $\bar{\lambda}(t)$ of $B(t)$ satisfies $\operatorname{Re} \bar{\lambda}(t) \leq -c_3$ or $\operatorname{Re} \bar{\lambda}(t) \geq c_4$

for some positive constants c_3 and c_4 is satisfied.

By (A-4) we are permitted to represent

$$B(t) = \operatorname{diag}(B_1(t), B_2(t)) \qquad (3.19)$$

where the block $B_j(t)$ have eigenvalues with real parts of the sign $(-1)^j$ throughout the sub-interval β_i, $i = 0,1,\ldots,k-1$. That means on each of the sub-intervals β_i we must get transformations $W_i(t)$ which transforms $B(t)$ to the form as in (3.19).

2.4 Decoupling the Difference Equations

In the preceding sections we have seen that Euler's Implicit method is $O(h^2)$ or $O(h)$ accurate for the fast variable y_2. The slow variable z_1 can also be solved by Euler's Implicit method provided that we take h_n smaller than the Lipschitz constant. This suggests to apply the scheme first and then decouple the variables if necessary.

In this section we study this approach of taking the difference equations and the possible decoupling transformation(s), applied on the difference equations.

So, consider the system

$$E\dot{x} = A(t)x + f(t) \qquad 0 < t < 1 \qquad (4.1)$$

with $E = \begin{bmatrix} I_m & 0 \\ 0 & \varepsilon I_k \end{bmatrix}$, $f = \begin{bmatrix} f^1 \\ f^2 \end{bmatrix}$

$x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$, $x^1 \in \mathbb{R}^m$, $x^2 \in \mathbb{R}^k$, similarly f^1, f^2;

$A(t) = \begin{bmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{bmatrix}$

Assume that (A-1) of § 2.2 holds. The discretization by Euler's Implicit scheme yields :

$$\frac{1}{h_n} E \begin{bmatrix} x^1_{n+1} - x^1_n \\ x^2_{n+1} - x^2_n \end{bmatrix} = \begin{bmatrix} A^{11}_{n+1} & A^{12}_{n+1} \\ A^{21}_{n+1} & A^{22}_{n+1} \end{bmatrix} \begin{bmatrix} x^1_{n+1} \\ x^2_{n+1} \end{bmatrix} + \begin{bmatrix} f^1_{n+1} \\ f^2_{n+1} \end{bmatrix} \tag{4.2}$$

(4.2) can also be written in the abbreviated form as :

$$(\frac{1}{h_n} E - A_{n+1}) x_{n+1} = \frac{1}{h_n} E\, x_n + f_{n+1} . \tag{4.3}$$

Introduce a sequence of transformations $\{T_n\}$ by defining (formally)

$$T_n = \begin{bmatrix} I_m & \varepsilon M_n \\ O & I_k \end{bmatrix} \tag{4.4a}$$

whose formal inverse is

$$T_n^{-1} = \begin{bmatrix} I_m & -\varepsilon M_n \\ O & I_k \end{bmatrix} \tag{4.4b}$$

As previously, put

$$z_n = T_n^{-1} x_n \tag{4.5}$$

to obtain (from (4.3))

$$E\, z_{n+1} = - E\, T_{n+1}^{-1}(h_n A_{n+1} - E)^{-1} E\, T_n z_n$$

$$- h_n E\, T_{n+1}^{-1}(h_n A_{n+1} - E)^{-1} f_{n+1} \quad (4.6)$$

But $h_n A_{n+1} - E = \begin{bmatrix} h_n A_{n+1}^{11} - I_m & h_n A_{n+1}^{12} \\ h_n A_{n+1}^{21} & h_n A_{n+1}^{22} - \varepsilon I_k \end{bmatrix}$

which implies that

$$(h_n A_{n+1} - E)^{-1} = \begin{bmatrix} C_3 & -C_4 \\ -D_3 & D_4 \end{bmatrix} \quad (4.7)$$

where,

$$C_3 = (h_n A_{n+1}^{11} - I_m - h_n A_{n+1}^{12}(h_n A_{n+1}^{22} - \varepsilon I_k)^{-1} h_n A_{n+1}^{21})^{-1}$$

$$D_4 = (h_n A_{n+1}^{22} - \varepsilon I_k - h A_{n+1}^{21}(h_n A_{n+1}^{11} - I_m)^{-1} h_n A_{n+1}^{12})^{-1}$$

$$C_4 = (h_n A_{n+1}^{11} - I_m)^{-1} h_n A_{n+1}^{12} D_4$$

$$D_3 = (h_n A_{n+1}^{22} - \varepsilon I_k)^{-1} h_n A_{n+1}^{21} C_3 .$$

Note that if the inverses above exist, then they are bounded. The inverses exist provided that $\min_n h_n |A^{ij}| < 1$. From (4.6) and (4.7), we find that

$$\begin{bmatrix} z^1_{n+1} \\ \\ \\ \varepsilon z^2_{n+1} \end{bmatrix} = \begin{bmatrix} -C_3 - \varepsilon M_{n+1} D_3 & -\varepsilon C_3 M_n - \varepsilon^2 M_{n+1} \cdot \\ & \cdot D_3 M_n + \varepsilon C_4 + \\ & + \varepsilon^2 M_{n+1} \quad D_4 \\ \varepsilon D_3 & \varepsilon^2 D_3 M_n - \varepsilon^2 D_4 \end{bmatrix} \begin{bmatrix} z^1_n \\ \\ \\ z^2_n \end{bmatrix}$$

$$+ \begin{bmatrix} C_3 + \varepsilon M_{n+1} D_3 & - C_4 - \varepsilon M_{n+1} D_4 \\ \\ -\varepsilon D_3 & D_4 \end{bmatrix} \begin{bmatrix} f^1_{n+1} \\ \\ f^2_{n+1} \end{bmatrix} \tag{4.8}$$

To claim that (4.8) is a (partial) decoupled form of (4.6), we must put the upper-right block zero, i.e.,

$$\varepsilon M_{n+1}(D_4 - D_3 M_n) - C_3 M_n + C_4 = 0. \tag{4.9}$$

This is a well-known Riccati algebraic equation. If we take M_{n+1} as the unknown, consequently, the Euler Implicit is applied forward, then (4.9) is an ill-conditioned equation, due to the appearance of ε with M_{n+1}. Instead, if we integrate the system with Euler's backward implicit scheme, then equation (4.9) might be interpreted as an equation for the unknown matrix M_n and (4.9) is no more ill-conditioned; moreover M_n remains bounded.

Matheij [74] considered a class of multistep methods to the system (4.1) with the stability criterion that $\mathrm{Re}\ \lambda(A^{22}(t)) < 0$, for all $t \in [0,1]$; he introduced such a sequence of transformations (4.4). The above interpretation of (4.9) reveals that one must use Euler's backward difference scheme, which is a member of the class of methods used in [74].

In the following we study the corresponding decoupling of the difference equations for the more usual decoupling (Lyapunov) transformations used, for example in Kokotovic et al. [62] (cf. § 2.2) This approach was also taken by O'Malley and Anderson [86] but for the constant coefficient case, i.e., when $A(t)$ does not depend upon t.

For this purpose, we take the difference approximation (4.3) and define a sequence $\{ S_n \}$ of transformations by

$$S_n = \begin{bmatrix} I_m & \varepsilon H_n \\ -L_n & I_k - \varepsilon L_n H_n \end{bmatrix} \qquad (4.10a)$$

where L_n and H_n are two matrices of dimensions $k \times m$ and $m \times k$ respectively to be defined (determined) later.

The formal inverse of S_n is :

$$S_n^{-1} = \begin{bmatrix} I_m - H_n L_n & -\varepsilon H_n \\ L_n & I_k \end{bmatrix} \qquad (4.10b)$$

Using the transformation (4.6) we put the difference equations (4.3) in the form below.

$$\text{Let} \quad z_n = S_n^{-1} x_n \qquad (4.11)$$

Then (4.3) admits the representation :

$$Ez_{n+1} = (ES_{n+1}^{-1}\bar{B}ES_n)z_n + (ES_{n+1}^{-1}\bar{B})f_{n+1} \qquad (4.12)$$

$$\text{where} \quad \bar{B} = (E - h_n A_{n+1})^{-1} = \begin{bmatrix} -C_3 & C_4 \\ D_3 & -D_4 \end{bmatrix} \qquad (4.13)$$

with C_3, C_4, D_3, D_4 as used in (4.7).

Using (4.13) in (4.12), we obtain

$$\begin{bmatrix} z_{n+1}^1 \\ \varepsilon z_{n+1}^2 \end{bmatrix} = \begin{bmatrix} C_1 & F_2 \\ F_1 & C_2 \end{bmatrix} \begin{bmatrix} z_n^1 \\ z_n^2 \end{bmatrix}$$

$$+ \begin{bmatrix} (-I_m + H_{n+1}L_{n+1})C_3 - \varepsilon H_{n+1}D_3 & (I_m - H_{n+1}L_{n+1})C_4 + \varepsilon H_{n+1}D_4 \\ -\varepsilon L_{n+1}C_3 + \varepsilon D_3 & \varepsilon L_{n+1}D_4 - \varepsilon D_4 \end{bmatrix} \begin{bmatrix} f_{n+1}^1 \\ f_{n+1}^2 \end{bmatrix}$$

(4.14)

where,

$$C_1 = -(I_m - \varepsilon H_{n+1}L_{n+1})C_3 - \varepsilon H_{n+1}D_3$$
$$-\varepsilon(I_m - \varepsilon H_{n+1}L_{n+1})C_4L_n + \varepsilon^2 H_{n+1}D_4L_n$$

$$C_2 = \varepsilon^2(-L_{n+1}C_3+D_3)H_n+\varepsilon^2(L_{n+1}C_4-D_4)(I_k-\varepsilon L_nH_n)$$

$$F_1 = -\varepsilon L_{n+1}C_3+\varepsilon D_3-\varepsilon^2(L_{n+1}C_4-D_4)L_n$$

$$F_2 = -\varepsilon(I_m-\varepsilon H_{n+1}L_{n+1})C_3H_n-\varepsilon^2H_{n+1}D_3H_n$$
$$+\varepsilon(I_m-\varepsilon H_{n+1}L_{n+1})C_4(I_k-\varepsilon L_nH_n)$$
$$+\varepsilon^2H_{n+1}D_4(I_k-\varepsilon L_nH_n).$$

Since the transformations of the type (4.6) contain two unknown matrices L and H, we can afford to decouple (4.3) completely. It is also necessary to do that since we need determining L as well as H. In (4.14) it is obvious that, for this purpose, we must put $F_1 = 0 = F_2$. Again $F_1 = 0$ simplifies the equation $F_2 = 0$ as well as the diagonal blocks C_1 and C_2. Hence, if L_n, L_{n+1} and H_n, H_{n+1} satisfy the equations :

$$L_{n+1}(C_3+\varepsilon C_4L_n) - (\varepsilon D_4L_n+D_3) = 0 \qquad (4.15a)$$

$$H_{n+1}(D_3-L_{n+1}C_3) - C_3L_n - (C_3+\varepsilon C_4L_n)H_nL_n = 0 \qquad (4.15b)$$

then the decoupled equations are :

$$z^1_{n+1} = -(C_3+\varepsilon C_4 L_n) z^1_n + g^1_{n+1} \qquad (4.16a)$$

$$\varepsilon z^2_{n+1} = (-D_4+\varepsilon L_n C_4) z^2_n + g^2_{n+1} \qquad (4.16b)$$

where $\begin{bmatrix} g^1_{n+1} \\ g^2_{n+1} \end{bmatrix}$ is the non-homogeneous term in (4.14).

Equation (4.15a) is a well known Riccati Algebraic equation which can be solved for L_{n+1} and leaves L_{n+1} bounded. Note that we start the recursion with $L_o = O$. (Matheij [74]). The same remarks apply to equation (4.15b), provided that we have already solved (4.15a). The recursion (4.15b) might also begin with $H_o = O$.

But if we take the Euler's implicit scheme backward, i.e., with the same scheme we want to integrate in backward direction, then it follows that we must solve the equations (4.15) for L_n and H_n. In that case, the matrix equations (4.15) become ill-conditioned. The solutions, thus can not be expected to be bounded.

2.5 Comments on Decouplings and Difference Schemes

In the preceding sections, we have seen that Lyapunov type transformation decouples the slow and fast variables completely, whereas Matheij type transformation decouples the slow variable from the system. For numerical purposes, however, the latter has advantages, since it needs only one Riccati Differential Equation to be solved, whereas the former needs two. This also holds for the approach taken in Section 2.4, where a difference scheme is first applied and the decoupling next. In this approach, Euler's Implicit scheme works well for the former transformation (a sequence of transformations) and Euler's Explicit scheme is appropriate for the latter (Matheij [74]). It is also important to know that in which direction the integration is to be carried out, backward or forward. Due to these restrictions, the methods have limited scopes, for example, they are appropriate for initial value problems. The direction of stability can also not be neglected (Theorem 1.1).

Therefore, we look at decoupling the states before application of a difference scheme. Decoupling might be done by any of the transformations :

Lyapunov type or the one introduced in § 2.2. However, we use the later for obvious computational convenience. Here again we need some clarifications as to how to solve the decoupled systems for given boundary conditions. If the boundary conditions are separated, i.e., they are given in separate sets for the fast and slow variables, then one might solve the decoupled slow variable first (e.g. by a centred difference scheme), and then use the results for the computation of the fast variable. If this is not the case, then one faces trouble as to generating an appropriate grid. For the slow variable, a grid $\{h_n\}$ is required, where h is less as compared to the Lipschitz constant of the slow system. On the other hand one might like to have h to be greater compared to the reciprocal of the norm of the system matrix of the fast system (see Theorem 3.1) to guarantee $e = O(h^2)$. This situation can be remedied by using a stretching (not necessarily uniform on the whole of $[0,1]$) of the independent variable. In fact, we need the stretching where the spectrum-structure changes. (see the Hybrid algorithm in § 3.2).

We have also remarked that the assumption that the eigenvalues with positive or negative real parts of the fast system need not be assumed to be separated

on the whole of [0,1]. (see (A-2)). Computationally, then we must determine the subintervals $[b_i, b_{i+1}]$ on which this happens, i.e., the structure of the spectrum does not change within any of the intervals (b_i, b_{i+1}). To this end, we select a partition b_i^1 , $0 = b_0^1 < b_1^1 \ldots < \ldots < b_{n_1}^1 = 1$; find the transformation $W(b_i^1)$ and block-diagonalize the system matrix of the fast system. This can be achieved by the usual QR-algorithm, or/and its modification by Kreiss et al. [68]. We check, whether the spectrum structure does change for different b_i^1's. If it changes for b_i^1 and b_{i+1}^1 , we might need another point in $[b_i^1, b_{i+1}^1]$. Note that, if the spectrum structure is not uniform throughout [0,1], then there are points b_i^j's, where the change of the structure is smooth, and our aim is to determine such points. (see §3.2 for details).

Our last remark concerns the global error. For the separated sets of boundary conditions, the global error can be estimated (see Dekker et al. [22]). But for the general boundary conditions, the global error, as such can not be estimated beforehand. Because, the structure of the spectrum is not known beforehand, and thus, the difference method is not known completely beforehand (except conditionally), condition number

of the error matrix can not be determined, and consequently the global error can't be estimated (see Kreiss et al. [68]).

Chapter 3

NUMERICAL DEMONSTRATIONS

3.1 Introduction

In the previous chapter we have analyzed the decoupling transformations and the difference scheme(s) for linear singularly perturbed systems. We have described the grid-generation and the numerical method that combines decoupling and difference schemes. The error analysis shows that the method is accurate upto $O(h^2)$ (or $O(h)$) at least locally. The global error analysis is not included, since, in general, it is not possible at this stage to do anything conclusively.

We realize that though a 'good theory' is a must, yet it does not conclude once and for all that a numerical method would never be unsuccessful. One loophole is the roundoff errors, which we do not include in the error analysis. Round-off errors depend heavily on the complexity of a method and the number of functional evaluations necessary for working it out. That is why it is instructive to do some numerical experimentations with any proposed method.

This chapter aims at fulfilling this gap between theory and practice. In § 3.2, the hybrid algorithm is presented, where we summarize the hybrid method in an algorithmic form to make implementation easier. In § 3.3, we take some examples to demonstrate the effectiveness of the method. The last section comments on the usefulness, effectiveness and the limitations of the algorithm.

3.2 Hybrid Algorithm

To facilitate understanding of the algorithm let us write down the problem we are considering and the underlying assumptions once again. We consider the linear system of singularly perturbed two-point boundary value problem :

$$\dot{y}_1 = A_{11}y_1 + A_{12}y_2 + f_1 \qquad (2.1a)$$

$$\varepsilon\dot{y}_2 = A_{21}y_1 + A_{22}y_2 + f_2 \qquad (2.1b)$$

for $0 < t < 1$ with the boundary conditions

$$B_o \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} + B_1 \begin{bmatrix} y_1(1) \\ y_2(1) \end{bmatrix} = B_2 \qquad (2.1c)$$

where '.' denotes d/dt, t - often called time,

ε is a small (compared to 1) positive parameter, $A_{ij}(t,\varepsilon)$ are smooth real-valued matrix functions of dimensions $n_i \times n_j$, f_j's are smooth real valued vector functions of dimensions n_j, where $n_j = m$, if $j = 1$ and $n_j = k$, if $j = 2$; B_0, B_1 are time-independent real matrices of dimensions $(m+k) \times (m+k)$; B_2 is an $m+k$ dimensional time-independent real vector so that (2.1c) represents $m+k$ linearly independent boundary conditions. We assume that :

(AL-1) There exists a partition $\{c_i\}$, $0 = c_0 < c_1 < \ldots < c_J = 1$ such that on each of the sub-intervals $[c_i, c_{i+1}]$ there exists a constant $K_i > 0$, so that the spectrum of $A_{22}(t,\varepsilon)$, $i = 0,\ldots,J-1$, can be represented by the union $S_- \cup S_+$ where $\mathrm{Re}\,\lambda < -K_i$, if $\lambda \in S_-$ and $\mathrm{Re}\,\lambda > K_i$, if $\lambda \in S_+$;

(AL-2) There are constants K_{ij}, $\bar{K}_{ij} > 0$ such that $||A_{ij}|| \leq K_{ij}$ and $||\dot{A}_{ij}|| \leq \bar{K}_{ij}$, $i,j = 1,2$.

Note that the smoothness of A_{ij} and f_j are included in the formulation of the problem, though, ofcourse, we only need them to be twice continuously differentiable.

An outline of the algorithm follows.

Our first task is to realize (AL-1) numerically i.e., to obtain the partition $\{ c_i \}$ and transform A_{22} to the block form D (see (2.11) of Ch. 2) on each of the sub-intervals $[c_i, c_{i+1}]$. We do it recursively. Given a partition point c_i, the next partition point c_{i+1} is determined as follows :

(A) If c_j, $j = 0,1,\ldots,i$ have been previously determined to be partition points, then compute the eigenvalues $\lambda_j(c_i+)$, where $\lambda_j(t) = \lambda_j(A_{22}(t))$, using say, QR-algorithm, thereby forming new sets of eigenvalues $S_-(c_i+)$, $S_+(c_i+)$ in addition to the former ones $S_-(c_i)$ and $S_+(c_i)$.

Note that $S(c_i+)$ means $\lim_{h \to 0+} S(c_i+h)$. Numerically c_i+ is to be interpreted as some point greater than c_i but sufficiently close to it. (Typically we take $c_i = c_i+\varepsilon/10$).

(B) Let $\bar{c} = c_i + \alpha$ be a trial value for c_{i+1}, where $\alpha = c_i - c_{i-1}$ (if $i = 0$, we take some trial value, typically = 1/16). Compute the eigenvalues $(\bar{c})$ and determine the sets $S_-(\bar{c})$ and $S_+(\bar{c})$.

(i) if no sets S change (compared to $S_-(c_i)$, and $S_+(c_i)$), go to (iv).

(ii) if the sets S change, mark $\bar{c}$ as a possible partition point. Before going to the next step, it is worthwhile to check, whether there is a partition point in between c_i and $\bar{c}$,

(iii) try (i), (ii) for $\alpha = (c_i - c_{i-1})/\sqrt{2}$. (It is advisable to try this again and again by taking smaller α's). It might also happen that we took α unnecessarily, i.e., the same structure of the spectrum is valid for a larger sub-interval.

(iv) if (iii) does not add anything new, try (i), (ii) for $\alpha = 2(c_i - c_{i-1})$. Go to (C).

The next task is to decouple the system in each of the sub-intervals $[c_i, c_{i+1}]$.

(C) Compute the matrices $W(t)$ and $D(t)$ using QR-algorithm (see Wilkinson [115] also Kreiss et al. i.e., lemmata : 9.1,9.2,[68]), on each of the sub-intervals $[c_i, c_{i+1}]$, i.e., on the grid-points. (see Remark 2.1).

Remark 2.1

We first take a uniform grid in each of the sub-intervals $[c_i, c_{i+1}]$ (typically grid length = 1/16). Then determine the matrices $W(t)$ and $D(t)$ on these grid-points; Note also that the grid should satisfy (3.6') of Ch. 2. Compute :

$$M_o = A_{12}A_{22}^{-1}W^{-1} \tag{2.2a}$$

$$M_1 = (A_{11}M_o - \dot{M}_o - M_oWA_{21}M_o - M_o\dot{W}W^{-1})D^{-1} \tag{2.2b}$$

$$M = M_o + \varepsilon M_1 \tag{2.2c}$$

$$B_{11} = A_{11} - MWA_{21} \tag{2.2d}$$

$$B_{22} = A_{22} + \varepsilon A_{21}MW \tag{2.2e}$$

$$g_1 = f_1 - MWf_2 \tag{2.2f}$$

Then the system (2.1a - 2.1b) is transformed into the form :

$$\dot{z}_1 = B_{11}z_1 + g_1 \tag{2.3a}$$

$$\varepsilon\dot{z}_2 = A_{21}z_1 + B_{22}z_2 + f_2 \tag{2.3b}$$

$$\text{with } y_1 = z_1 + MWz_2 \ , \ z_2 = y_2 \tag{2.3c}$$

on each of the sub-intervals $[c_i, c_{i+1}]$. Our task now, is to solve the system (2.3a - 2.3b) with (2.1c).

Note that the boundary conditions (2.1c) should also be transformed as well. This is done using (2.3c). Write the transformed boundary conditions as :

$$C_o \begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} + C_1 \begin{bmatrix} z_1(1) \\ z_2(1) \end{bmatrix} = C_2 \quad . \qquad (2.4)$$

Note that the system ((2.3a) -(2.3b)) is an m+k dimensional system, with (2.3a) being of dimension m and (2.3b) can be represented as two systems with dimensions k_{1i}, k_{2i} ($k_{1i} + k_{2i} = k$) on each of the sub-intervals $[c_i, c_{i+1}]$, k_{1i} being the order of D_1 and k_{2i}, that of D_2.

We intend to solve the slow system (2.3a) by Trapezoidal rule and the fast system (2.3b) by the Euler's (backward and forward) schemes. Hence according to the above paragraph, trapezoidal rule is applied for the first m dimensional system, Euler's forward Implicit scheme for the next k_{1i} dimensional system and Euler's backward Explicit scheme (see § 2.5) for the last k_{2i} dimensional system. This is done in (D).

Remark 2.2

For the methods to be accurate, the grid must be generated accurately, i.e., the grid $\{h_n\}$ must satisfy : (see (3.6') of Ch. 2)

$$\frac{2\varepsilon}{h_n} < ||B_{22}|| < \frac{h_n}{\varepsilon}$$

$$h_n < ||B_{11}|| \qquad \text{for all } n.$$

If possible, the condition that $h_n|B_{22}(t_{n+1})| > 1$ also be verified. A sub-interval where the grids are unusually refined compared to other parts will be referred to as a stretching sub-interval, and its end-points as stretching points.

(D) Denote

$$\delta_{ij} = \begin{cases} \frac{1}{2}, & j = 1,2,\ldots,m \\ 0, & j = m+1,\ldots,m+k_{1i} \\ 1, & j = m+k_{1i}+1,\ldots,m+k. \end{cases}$$

(The first subscript signifies the choice of δ_{ij} is dependent upon the sub-intervals $=c_i, c_{i+1}]$.)

Denote $\Delta_i = \text{diag}\,(\delta_{ij})$. The (hybrid) difference scheme is :

$$\frac{u_{n+1} - u_n}{h_n} = \Delta_i B_n u_n + (I - \Delta_i) B_{n+1} u_{n+1} + \Delta_i G_n + (I - \Delta_i) G_{n+1}$$

$$\text{for } t_n \in [c_i, c_{i+1}]. \quad (2.5a)$$

where, we have abbreviated (2.3a - 2.3b) by taking

$$B = \begin{bmatrix} B_{11} & 0 \\ \varepsilon^{-1} A_{21} & \varepsilon^{-1} B_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_1 \\ f_2 \end{bmatrix} \quad (2.5b)$$

and the usual notation x_n means (x : u, B, G etc.)

$$x(t_n), \qquad t_n = \sum_{i=o}^{n-1} h_n, \qquad n = 0, \ldots, N$$

The boundary conditions are added with the difference system as usual by taking

$$u_o = \begin{bmatrix} z_1(a) \\ z_2(a) \end{bmatrix}, \qquad u_N = \begin{bmatrix} z_1(b) \\ z_2(b) \end{bmatrix}$$

(for details, see Keller [55]).

(E) The numerical solution at the points not included in the grid is determined by using Richardson's extrapolation [1].

Remark 2.3

Note that stretching of the boundary layers

etc. is done automatically in the form of grid-generation due to the conditions in Remark 2.2 that are to be monitored in the grid-generation process.

Remark 2.4

An alternative approach is to solve Equation 2.10 of Ch. 2 approximately by using asymptotic series expansions. For example $O(\varepsilon^2)$ approximation would look like : $L = L_o + \varepsilon L_1 + O(\varepsilon^2)$ with $L_o = A_{12}A_{22}^{-1}$

$L_1 = ((A_{11} - L_oA_{21})L_o - \dot{L}_o)A_{22}^{-1}$. This approximation of L lets the system matrix of the decoupled equations look like

$$\begin{bmatrix} A_{11} - LA_{21} & 0 \\ A_{21} & \varepsilon A_{21}L + A_{22} \end{bmatrix} .$$

Note that this is to be done on the whole of $[a,b]$ at once (contrary to the previous approach, where $W(t)$ was involved.] Then we find out the transformations $W(t)$ on the sub-intervals $[c_i, c_{i+1}]$.

This approach is justified due to Lemma 3.1 in Chapter 2.

3.3 Examples

We have seen the difficulties associated with linear systems of singularly perturbed two-point boundary-value problems in § 1.1. The difficulties were, in general, two fold from the view point of computations. As was noted there, computation of the solution should take care of the situations with large eigenvalues and situations with the largely varying eigenvalues. Again, asymptotic analysis of these problems reveals the fact that, very often, in the practical problems, there occur boundary layers near the ends $t = 0$ and $t = 1$ of the interval of integration $[0,1]$.

The corner layers or the interior layers are also observed to occur in the middle of the interval.

In this section, we consider some examples where we meet one or many of the aforesaid situations. Figures following the examples show the numerical solutions. They intend to show that in the choosen grid the solution do not contain any oscillatory elements as well as they give ideas of how a solution looks like. Only for modest ε's, these figures are drawn. The tables following the examples

provide the absolute value of the difference between the exact/asymptotic solution and the numerical solution obtained by the hybrid method. In the tables the following abbreviations are used.

SP : Stretching points (end points of SS)

SS : Stretching sub-intervals, where the grids are made finer.

NG : Number of grid-divisions used in the corresponding stretching sub-interval.

Mad : Maximum of the absolute values of the differences between the exact/asymptotic solution and the numerical solution.

Mid : Minimum of the absolute values of the differences between the exact/asymptotic solution and the numerical solution.

For each of the problems, we give two tables; where the first table gives the choice of stretching points and the second, showing the (comparison of) results in terms of Mad. and Mid.

From amongst many examples experimented, the following model example is choosen from O'Malley [82] for illustration purposes. The parameters α, β in the following example, chosen appropriately, give

rise to different nature of the solution. We will also come across the difficulties associated with the eigenvalues of the system matrix, when the different cases come into picture.

The model linear problem, with the constants (to be fixed later) α, β is :

$$\varepsilon \ddot{y} + 2\alpha t\dot{y} - \alpha\beta y = 0 \qquad -1 < t < 1 \qquad (3.1a)$$

with the prescribed boundary conditions given at $t = \pm 1$. (3.1b)

The general solution of (3.1) is

$$y(t) = \exp\left(-\frac{\alpha t^2}{\varepsilon}\right)\left[c_1 D_{-1} - \frac{1}{2}\beta \left(\frac{2\alpha}{\varepsilon}\right)^{\frac{1}{2}} t\right) + c_2 D_{\frac{1}{2}\beta}\left(i\left(\frac{2\alpha}{\varepsilon}\right)^{\frac{1}{2}} t\right)] \qquad (3.2)$$

for arbitrary constants c_1, c_2. The constants c_1, c_2 might be determined from the boundary conditions, i.e., from

$$y(\pm 1) = \exp\left(-\frac{\alpha}{2\varepsilon}\right)\left[c_1 D_{-1 - \frac{1}{2}\beta}\left(\left(\pm \frac{2\alpha}{\varepsilon}\right)^{\frac{1}{2}}\right) + c_2 D_{\frac{1}{2}\beta}\left(\pm i\left(\frac{2\alpha}{\varepsilon}\right)^{\frac{1}{2}}\right)\right]. \qquad (3.3)$$

The asymptotic approximation of D_n's are :

$$D_n(z) = \exp\left(-\frac{z^2}{4}\right) z^n (1 + o(1)$$

as $|z| \to \infty$, for $n = 0,1,2,\ldots$

which might be needed in solving (3.3) for c_1, c_2.

The asymptotic behaviour of (3.2) differ according to the values of the constants α and β. We have the following four cases accordingly.

Case 1

$$\alpha > 0, \ \beta \neq -2n, \ n = 1,2,\ldots .$$

In this case the solution $y(t)$ in (3.2) can be represented (asymptotically) as :

for $-1 \leq t < 0$,

$$y(t) = (-t)^{\frac{\beta}{2}}[\ y(-1) - (-1)^{\frac{\beta}{2}} y(1) + o(1)] + ((-1)^{\frac{\beta}{2}} y(1) + o(1)$$

for $0 < t \leq 1$,

$$y(t) \sim \frac{\Gamma(-1 - \frac{\beta}{2})}{2(\frac{\pi\alpha}{\varepsilon})^{\frac{1}{2}}} (y(-1) - (-1)^{\frac{\beta}{2}} y(1) + o(1)) \times \frac{\exp(-\alpha t^2/\varepsilon)}{t^{1+\beta/2}} + ((-1)^{\frac{\beta}{2}} y(1) + o(1))(-t)^{\frac{\beta}{2}}$$

and $y(0) = O(\varepsilon^{\beta/4})$; $\Gamma(.)$ being the gamma function. This represents non-uniform behaviour at $t = 0$, an interior layer near $t = 0$.

Case 2

$$\alpha < 0, \quad \beta \neq 2m, \quad m = 0,1,2,\ldots$$

The solution is :

for $-1 \leq t < 0$,

$$y(t) = \frac{1}{t^{1+\beta/2}} (y(-1) + o(1)) \exp(\alpha(1-t^2)/\varepsilon) +$$

$$+ O(e^{\alpha(1-\delta)/\varepsilon})$$

for any $\delta > 0$.

for $0 < t \leq 1$,

$$y(t) = t^{-(1+\frac{\beta}{2})} (y(1) + o(1)) \exp(\alpha(1-t^2)/ \quad .$$

and $y(0) = O(e^{\alpha(1-)/})$

for any > 0.

This shows the non-uniform behaviour of the solution at both of the end-points $t = \pm 1$.

Case 3

$$\alpha < 0, \quad \beta = 2m, \quad m = 0,1,2,\ldots$$

The solution in this case is :

for $t \in [-1,0) \cup (0,1]$,

$$y(t) = \frac{1}{2}[(y(1) - (-1)^m y(-1)) \frac{\exp(\alpha(1-t^2)/\varepsilon)}{t^{1+m}}$$

$$+ (y(1) + (-1)^m y(-1))t^m] + o(1)$$

Case 2. $\alpha < 0, \quad \beta \neq 0,2,4, \ldots$

Limiting solution as $\varepsilon \to 0$ vanishes on $-1 < t < 1$.

Case 3. $\alpha < 0, \quad \beta = 0,2,4, \ldots$

Limiting solution as $\varepsilon \to 0$ is :

$$\frac{1}{2}[y(1)t^{\beta/2} + y(-1)(-t)^{\beta/2}] \quad \text{on} \quad -1 < t < 1.$$

Case 4. $\alpha > 0, \quad \beta = -2, -4, \ldots$

Solution becomes exponentially large on $(-1,1)$.

For computational purposes, we choose α, β, $y(-1)$ and $y(1)$ appropriate to each of the above four cases. In the following, the Example 1.m correspond to the case - m.
Note that a problem is always transformed to a system in the standard way.

Example 1.1 $\alpha = 1, \quad \beta = 2.$ (Case 1)

Explicitly, the problem is :

$$\varepsilon\ddot{y} + 2t\dot{y} - 2y = 0 \qquad -1 < t < 1$$
$$y(-1) = -1, \quad y(1) = 2.$$

Example 1.2 $\alpha = -1, \quad \beta = 1.$ (Case 2)

i.e.,

$$\varepsilon\ddot{y} - 2t\dot{y} + y = 0 \qquad -1 < t < 1$$
$$y(-1) = -1, \quad y(1) = 2.$$

Example 1.3 $\alpha = -1, \quad \beta = 0$ (Case 3)

i.e.,

$$\varepsilon \ddot{y} - 2t\dot{y} = 0 \quad , \qquad -1 < t < 1$$

$$y(-1) = -1, \; y(1) = 2.$$

Example 1.4 $\alpha = 1, \quad \beta = -2$ (Case 4)

i.e.,

$$\varepsilon \ddot{y} + 2t\dot{y} + 2y = 0 \qquad -1 < t < 1$$

$$y(-1) = -1, \; y(1)=2$$

Note that the solution to Example 1.4 has exponential growth and damping in the left and right boundary layers respectively; and it is unbounded elsewhere. Due to this reason, we do not have graphical solution for Example 1.4. Figures (3.1 - 3.3) show the numerical solutions for the Examples (1.1) - (1.3) with $\varepsilon = 10^{-4}$. Table (3.1 - 3.4) summarize the results of comparison of the numerical solutions with the asymptotic solutions.

Note that the interval of integration is $[-1,1]$ and not $[0,1]$ which we adopted in discussing the hybrid method. This is handled by first transforming the interval $[-1,1]$ to $[0,1]$ before the application of the hybrid method. Alternatively, in the algorithm itself, any interval of the type $[a,b]$ can

be retained asking the user to provide the values [a,b], and translating this[0,1] in the whole of the hybrid method into [a,b].

To demonstrate the effectiveness of the hybrid method for higher-dimensional problems, the following example is chosen. This arose from the eigenvalue problem in vibration of strings (Lord Rayleigh [71]). However, we will fix the eigenvalue

$$\lambda = \pi^2 + 4\varepsilon\pi^2(+ O(\varepsilon^2)) .$$

in accordance with O'Malley [82, § 3.3].

Example 2

Consider the two-point boundary-value problem :

$$2\frac{d^4y}{dt^4} - \frac{d^2y}{dt^2} = \lambda y \qquad 0 < t < 1 \qquad (3.4a)$$

with the boundary conditions

$$y(0) = \dot{y}(0) = y(1) = \dot{y}(1) = 0 \qquad (3.4b)$$

where,

$$\lambda = \pi^2 + 4\varepsilon\pi^2. \qquad (3.4c)$$

The asymptotic solution for (3.4) is given by :

$$y(t,\varepsilon) = \sqrt{2}\,\pi\,\Big[\frac{\sin \pi t}{\pi} + \varepsilon\Big\{\frac{\sin \pi t}{\pi} - (1-2t)\cos \pi t + e^{-t/\varepsilon} - e^{-(1-t)/\varepsilon}\Big\} + O(\varepsilon^2)\Big]. \tag{3.5}$$

The numerical solution for (3.4) with $\varepsilon = 10^{-6}$ is shown in Figure 3.4 and the results of comparison of the asymptotic solution (3.5) with the numerical solution obtained by the hybrid method are summarized in Table 3.5.

The experiences with the numerical experimentation are presented in § 3.4.

Table 3.1a

Choice of Stretching Points for Example 1.1

ε | SP	b_o	b_1	b_2	b_3
1E-2	-1	-0.05	0.05	1
1E-4	-1	-0.003	0.004	1
1E-6	-1	-0.0018	0.0008	1

Table 3.1b

Summary of the Comparison Results for Example 1.1

ε	SS	$[b_o,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$
	NG	12	20	8
1E-2	Mad	24E-3	88E-3	28E-4
	Mid	OE-1	1E-6	OE-1
	NG	16	16	12
1E-4	Mad	22E-5	7E-5	4E-4
	Mid	1E-7	1E-8	OE-1
	NG	16	20	12
1E-6	Mad	2E-5	22E-6	1E-5
	Mid	OE-1	OE-1	OE-1

Table 3.2a

Choice of Stretching Points for Example 1.2

ε | SP	b_o	b_1	b_2	b_3
1E-2	-1	-0.96	0.964	1
1E-4	-1	-0.992	0.993	1
1E-6	-1	-0.991	0.9992	1

Table 3.2b

Summary of the Comparison Results for Example 1.2

ε	SS	$[b_o,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$
	NG	12	16	8
1E-2	Mad	2E-2	2E-2	21E-3
	Mid	OE-1	1E-9	OE-1
	NG	12	20	12
1E-4	Mad	22E-5	29E-5	16E-5
	Mid	OE-1	4E-6	OE-1
	NG	12	24	12
1E-6	Mad	15E-5	14E-5	12E-5
	Mid	1E-9	1E-8	1E-10

Table 3.3a

Choice of Stretching Points for Example 1.3

ε	SP	b_o	b_1	b_2	b_3
1E-2		-1	-0.98	0.976	1
1E-4		-1	-0.992	0.998	1
1E-6		-1	-0.9993	0.999	1

Table 3.3b

Summary of the Comparison Results for Example 1.3

ε	SS	$[b_o,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$
	NG	8	16	12
1E-2	Mad	14E-3	20E-3	28E-3
	Mid	OE-1	4E-6	OE-1
	NG	16	16	16
1E-4	Mad	25E-5	35E-5	8E-4
	Mid	OE-1	2E-6	OE-1
	NG	16	24	16
1E-6	Mad	33E-6	19E-6	9E-5
	Mid	OE-1	1E-7	OE-1

Table 3.4a

Choice of Stretching Points for Example 1.4

ε	SP	b_0	b_1	b_2	b_3	b_4	b_5
1E-1		-1	-0.89	-0.01	0.01	0.91	1
1E-2		-1	-0.849	-0.007	0.006	0.87	1
1E-3		-1	-0.84	-0.006	0.006	0.84	1

Table 3.4b

Summary of the Comparison Results for Example 1.4

ε	SS	$[b_0,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$	$[b_3,b_4]$	$[b_4,b_5]$
	NG	8	8	8	8	8
1E-1	Mad	24E-3	*	*	*	21E-2
	Mid	1E-12	*	*	*	22E-7
	NG	12	8	8	8	12
1E-2	Mad	7E-4	*	*	*	11E-4
	Mid	OE-1	*	*	*	12E-7
	NG	16	8	12	8	12
1E-3	Mad	12E-5	*	*	*	9E-5
	Mid	OE-1	*	*	*	OE-1

*indicates overflow in the numerical solution.

Table 3.5a

Choice of Stretching Points for Example 2

ε	SP	b_o	b_1	b_2	b_3
1E-2		0	0.06	0.95	1
1E-4		0	0.005	0.994	1
1E-5		0	0.0002	0.998	1

Table 3.5b

Summary of the Comparison Results for Example 2

ε	SS	$[b_o,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$
	NG	8	16	8
1E-2	Mad	25E-3	9E-2	3E-3
	Mid	OE-1	2E-6	OE-1
	NG	16	16	8
1E-4	Mad	28E-5	32E-5	4E-4
	Mid	OE-1	3E-6	OE-1
	NG	12	20	8
1E-5	Mad	31E-6	25E-6	2E-5
	Mid	OE-1	5E-7	OE-1

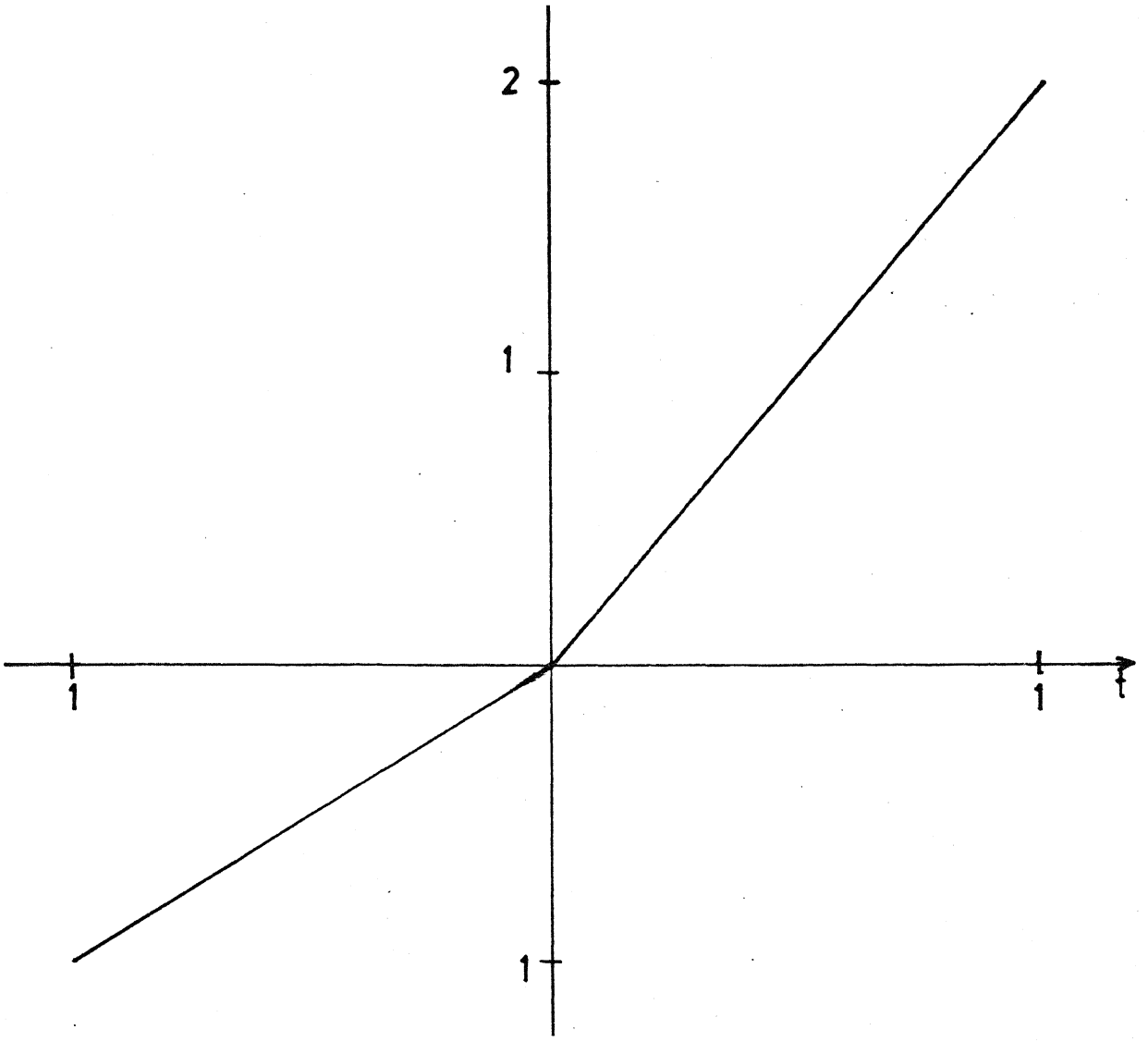


Figure 3.1

Numerical solution to example 1.1, $\varepsilon = 10^{-4}$

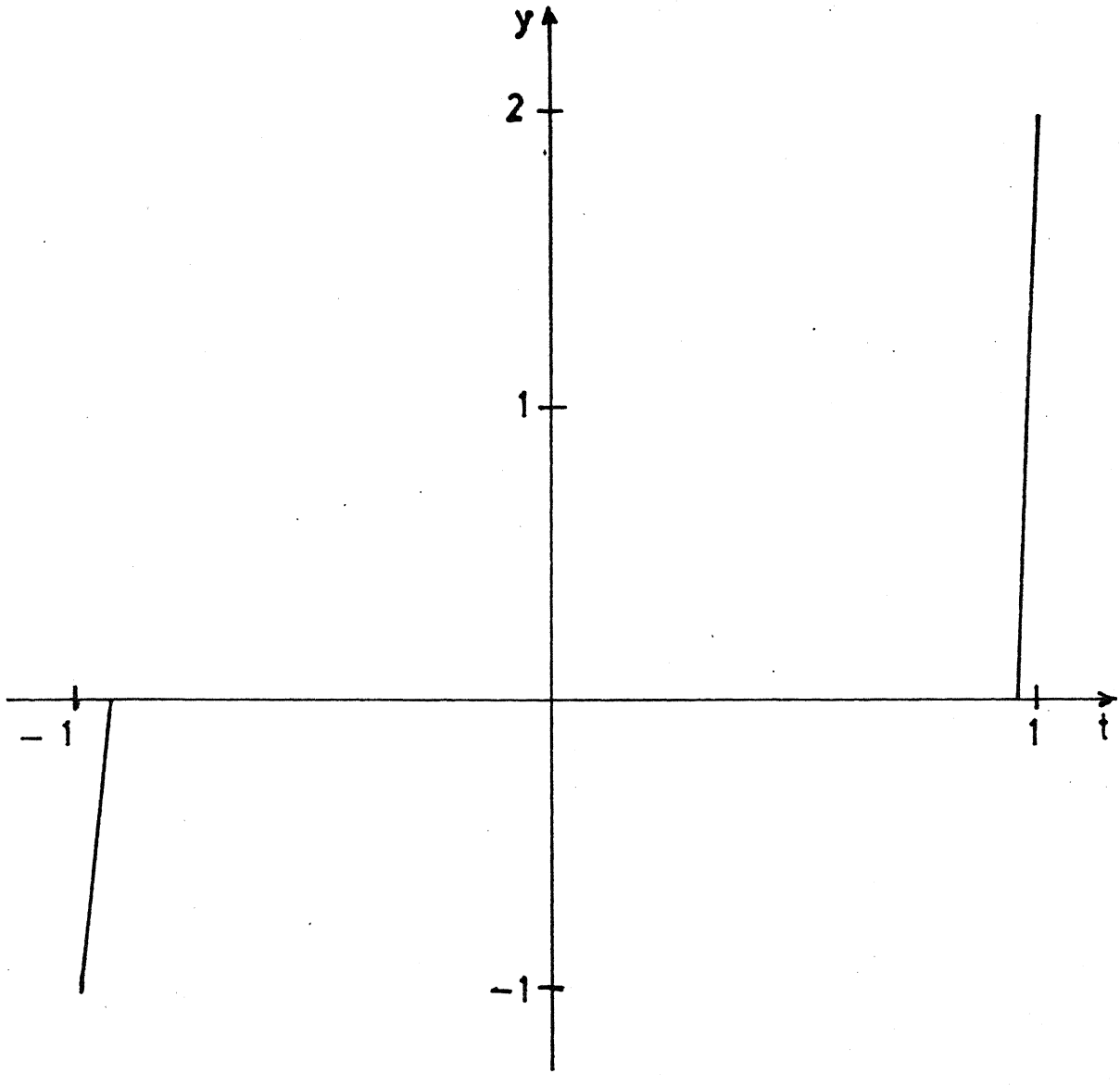


Figure 3.2

Numerical solution to example 1.2 , $\varepsilon = 10^{-4}$

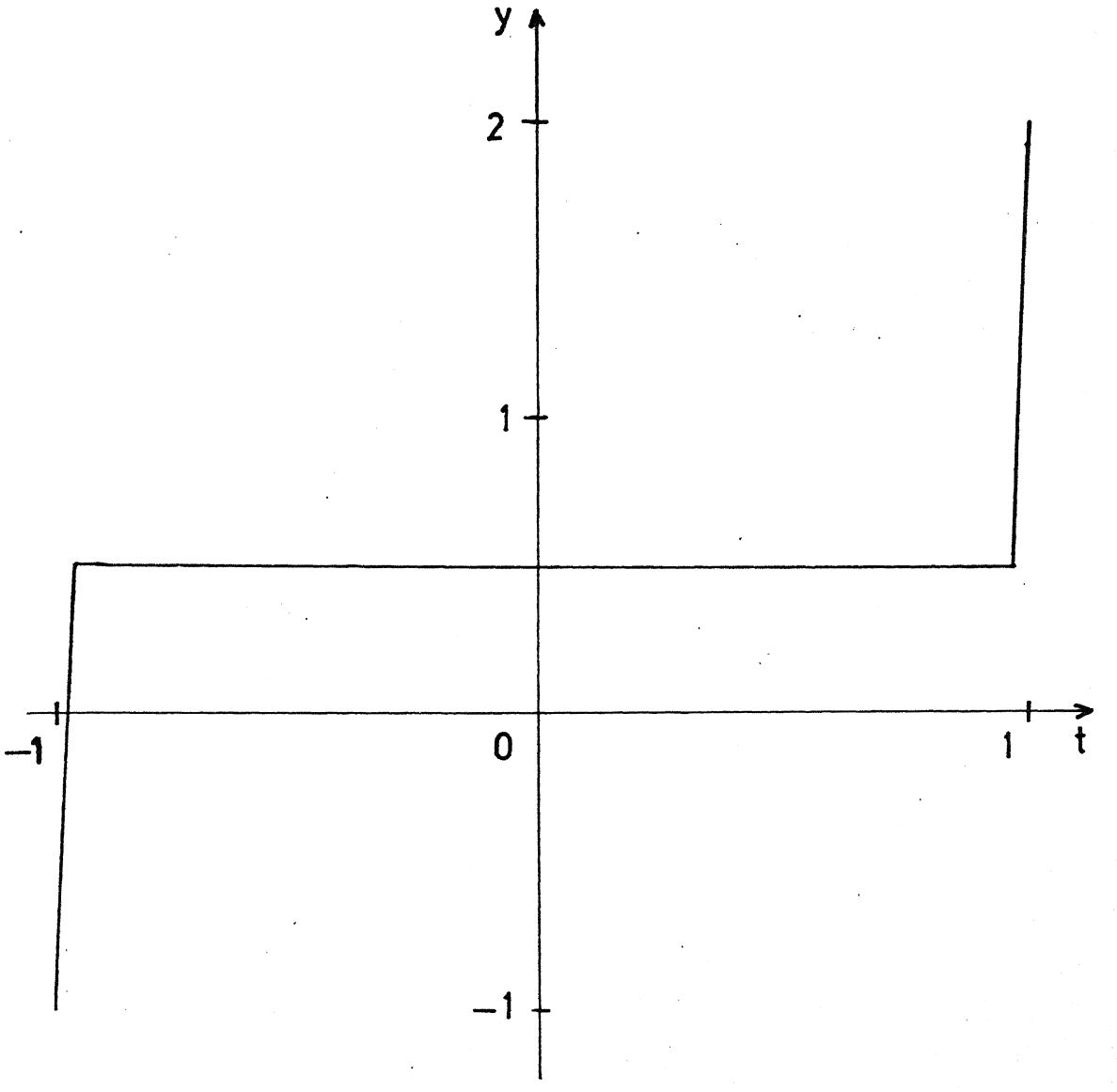


Figure 3.3

Numerical solution to example 1.3, $\varepsilon = 10^{-4}$

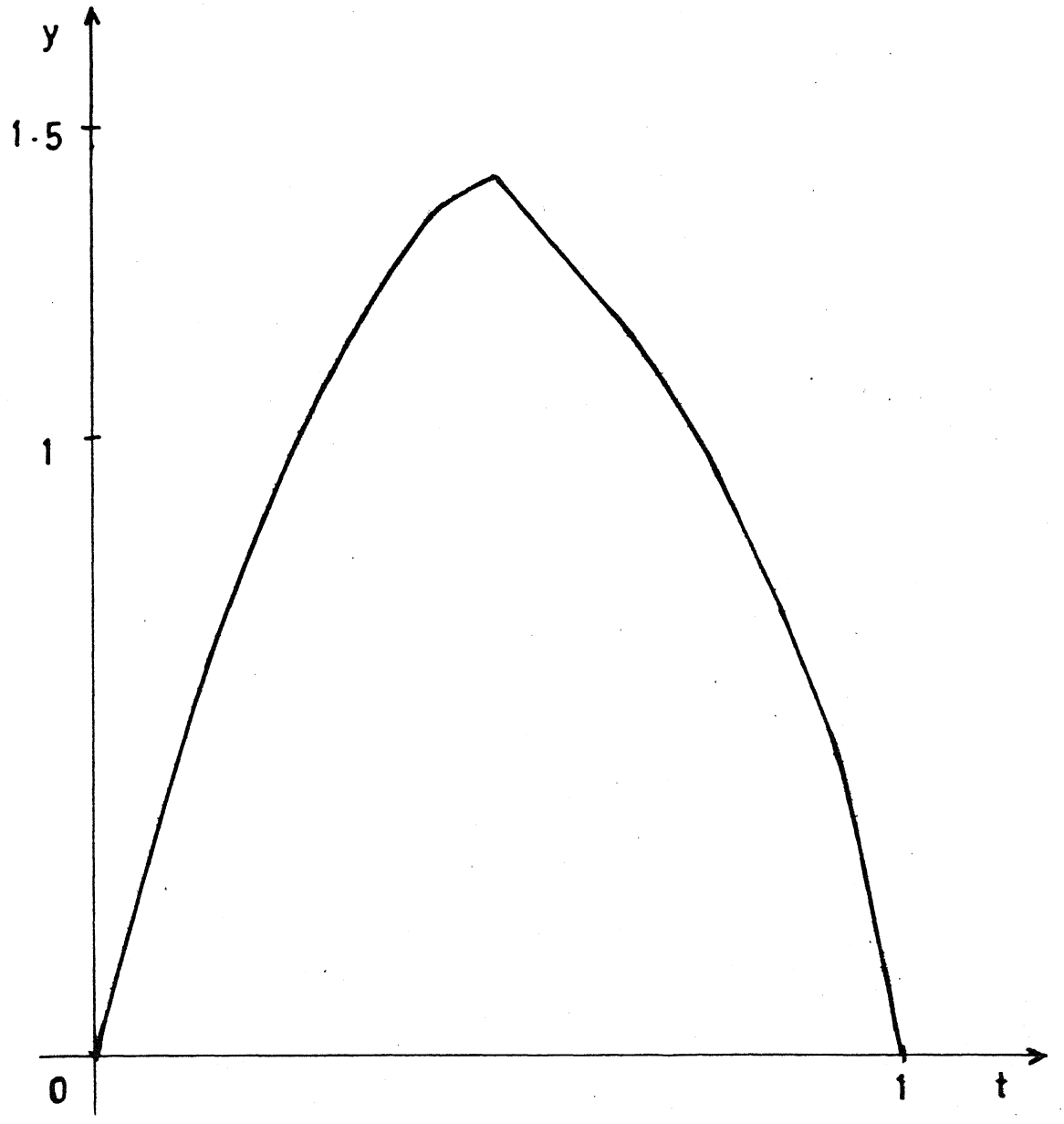


Figure 3.4

Numerical solution to example 2 , $\varepsilon = 10^{-6}$

3.4 Discussions

Examples (1.2) - (1.3) exhibit boundary layers near both the end points $t = \pm 1$. The asymptotic solutions are constructed upto $O(\varepsilon^2)$ accuracy, which can be accepted as fairly well approximations of the solutions to these examples. From the results in Tables 3.2 - 3.3 , we see that the numerical solution compares well with the asymptotic solution. Again, as ε decreases, the numerical solution compares better with the asymptotic solution.

Example 1.1 demonstrates the effectiveness of the hybrid method for interior layers. Here also as ε decreases, the numerical approximation compares well with the asymptotic solution.

In case of Example 1.4 it appears that the method does not work properly. But a more involved thought reveals the fact that the general solution

$$y(t) = (c_1 + c_2 \int e^{\frac{t^2}{\varepsilon}} dt)\, e^{-\frac{t^2}{\varepsilon}}$$

of the problem remains unbounded as ε approaches zero. And in the numerical solution, we see that for smaller ε's, overflow occurs; which is exactly what we expect. It is also observed that whenever,

there is overflow in the numerical solution, the asymptotic solution is also too large. This demonstrates the efficiency of the numerical method. This also suggests that much of the qualitative natures of the solution might well be known by solving problems with the present numerical method.

Example 2 illustrates the effectiveness of the hybrid method for higher dimensional problems.

The figures indicate the absence of small oscillations in the numerical solution. In fact, the grid-generation helps to resolve the possible oscillations. It should also be noted that we have restricted the trapezoidal rule to the slow component which could otherwise produce the oscillations. The numerical experimentation also tallies with the expected robustness of the method.

In general, the thickness of the boundary layers or the interior layers are not known, analytically or asymptotically. But the choice of the stretching points (partition points) in the method indicates the numerical thickness of the layers. This provides us a necessary information, which could not be known otherwise. Unfortunately, this is not conclusive, since the grid-generation

only chooses the stretching points by looking for the 'possibility' of the existence of the layers and not 'necessity'. However, after one gets the numerical solution, one can check, whether there is a rapid change of the solution in the concerned region; which would perhaps be conclusive for locating the layers and determining their thickness(es) numerically.

Chapter 4

SPECIAL METHODS

4.1 Introduction

As demonstrated in Chapter 3, the hybrid method solves linear systems of singularly perturbed two-point boundary-value problems accurately. The hybrid method is a general purpose method; and naturally it suffers from the demerits common to all the general purpose methods. The demerit is its lack of simplicity in computation. For simple homogeneous problems, without turning points and interior layers, the same procedure can be applied. Suppose that, from physical considerations, one knows that the problem at hand contains only the possible boundary layers, and for certainty there are no turning points or interior layers. So one naturally feels reluctant to proceed through the time-consuming grid-generation process and decoupling. This is the beginning stage of looking for the special methods which would avoid such unnecessary computation as well as save machine time.

This chapter presents some special numerical techniques which avoid the cumbersome grid-generation

and compute the numerical solutions directly. Note that the extra information the hybrid method was providing us is the layer-thickness. Even for a sub-class of problems such as homogeneous equations possibly with only boundary layers, for which we will apply the special methods, there is no way of determining the layer-thicknesses. This is rather a drawback of the special methods going to be presented in this chapter. However, using asymptotic analysis and its procedure of solution, we can get a fairly good idea of this thickness, which we present in § 4.2. In § 4.3 and § 4.4 two methods are suggested and also illustrated by means of some examples. The last section, § 4.5 discusses the effectiveness and scope of applications of these special methods.

4.2 Boundary Layer Thickness

Consider the following singularly perturbed system of ordinary differential equations :

$$\dot{x} = A(t,\varepsilon)x + B(t,\varepsilon)y \qquad (2.1a)$$

$$0 < t < 1$$

$$\varepsilon\dot{y} = C(t,\varepsilon)x + D(t,\varepsilon)y \qquad (2.1b)$$

with

$x_1(0)$, $x_2(1)$, $y_1(0)$, $y_2(1)$ prescribed; (2.2)

where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$,

$x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, $y_1 \in \mathbb{R}^{n_3}$, $y_2 \in \mathbb{R}^{n_4}$, and '.' denotes d/dt, A, B, C, D are matrices of functions of t and ε with appropriate dimensions; and ε is a small positive parameter ($0 < \varepsilon << 1$). Problems of this type have been well studied via asymptotic expansion techniques [41, 112]. Under the principal assumption that the eigenvalues of the $(n_3+n_4) \times (n_3+n_4)$ matrix $D(t,0)$ have non-zero real parts, Harris [41] shows that such systems have an n_1+n_2 dimensional manifold of solutions which tend to the solutions of the reduced system (i.e., the system obtained from (1) by taking $\varepsilon = 0$) as $\varepsilon \to 0$. Moreover, if $D(t,0)$ has k, $0 \leq k \leq n_3+n_4$ stable eigenvalues, then there are k linearly independent solutions of (2.1) which decay to zero (like $e^{-M_1 t/\varepsilon}$, for some positive definite matrix M_1) as the stretched variable $\tau = t/\varepsilon \to \infty$, and there are n_3+n_4-k linearly independent solutions which decay to zero (like $e^{-M_2(1-t)/\varepsilon}$ for some positive definite matrix M_2) as the stretched

variable $\sigma = (1-t)/\varepsilon \to \infty$. Since the general solution of (2.1) is a linear combination of these (asymptotic) solutions, any solution of (2.1) satisfies a three time-scale property. That is, such an asymptotic solution needs to be an additive sum of functions of t, τ and σ where functions depending upon τ and σ represent the boundary-layer behaviours of the solutions at $t = 0$ or $t = 1$ and function dependent upon t satisfies the reduced system corresponding to (2.1).

The question is exactly when one can neglect the functions of τ and σ to represent the solution of (2.1) only by the reduced solution. In the literature, these regions beyond which the reduced solution represents the solution of (2.1) are termed as boundary layers and their measure, as the boundary layer thickness(es). An inequality concerning this thicknesses was given in Hsiao and Jordan[46] for scalar equations of the type

$$\varepsilon\ddot{u} + p(t)\dot{u} + q(t)u = f(t) \qquad 0 < t < 1 \qquad (2.3)$$

with $u(0)$, $u(1)$ prescribed.

In this section, an estimate of this thickness for systems of the type (2.1) is tried; and an inequality

concerning the norm of $\exp(D(0,0)th)$ is given, 'th' being the thickness of the boundary layer(s).

For our purpose, the following theorem will be sufficient [16,41,83].

Theorem 2.1

Assume that

(H-1) All eigenvalues of the matrix $D(t,0)$ have non-zero real parts on $[0,1]$,

(H-2) $D(t,0)$ is non-singular for $t \in [0,1]$,

(H-3) The reduced system

$$\dot{X} = (A(t,0) - B(t,0)D^{-1}(t,0)C(t,0))X \quad (2.4a)$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad X_1(0) = x_1(0), \quad X_2(1) = x_2(1) \quad (2.4b)$$

has a unique solution.

(H-4) The initial boundary layer problem

$$\frac{d}{d\tau}\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = D(0,0)\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad (2.5)$$

has a decaying solution on $0 \leq \tau < \infty$ for any initial value $U_1(0)$; $U_1 \in \mathbb{R}^{n_3}$, $U_2 \in \mathbb{R}^{n_4}$;

(H-5) The terminal boundary layer problem

$$\frac{d}{d\sigma}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = -D(1,0)\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (2.6)$$

has a unique decaying solution on $0 \leq \sigma < \infty$ for any initial value $V_2(0)$; $V_1 \in \mathbb{R}^{n_3}$, $V_2 \in \mathbb{R}^{n_4}$.

Then, the system (2.1 - 2.2) has a unique solution of the form :

$$x(t,\varepsilon) = X(t) + O(\varepsilon) \tag{2.7a}$$

$$y(t,\varepsilon) = Y(t) + U(\tau) + V(\sigma) + O(\varepsilon) \tag{2.7b}$$

on $0 \leq t \leq 1$, where the terms X, Y, U, V all have asymptotic power series expansions and the functions of $\tau = t/\varepsilon$ and $\sigma = (1-t)/\varepsilon$ decay to zero as τ , $\sigma \to \infty$.

Remark 2.1

Instead of assuming the existence-uniqueness of the solutions of (2.4 - 2.6) explicitly as in (H-2) - (H-5). We could have taken some sufficient conditions to guarantee these [16].

The end-conditions to be supplied with (2.5) and (2.6) for getting a solution of (2.1) are

$$U_1(0) = y_1(0) - Y_1(0), \; U_2(\infty) = 0 \tag{2.8}$$

and

$$V_2(0) = y_2(1) - Y_2(1), \; V_1(\infty) = 0. \tag{2.9}$$

But, since the solutions of boundary-layer problems (2.5) and (2.8), and (2.6) and (2.9) become insignificant outsidide the boundary layer regions (due to their exponential decaying nature), it is appropriate to replace the conditions $U_2(\infty) = 0$ and $V_1(\infty) = 0$ by $U_2(\tau_1) = 0$ and $V_1(\sigma_2) = 0$ for some τ_1 , $\sigma_2 < \infty$, respectively [109]. We require the solutions of the modified problems (taking τ_1 or σ_2 in place of ∞) to differ from that of the unmodified ones only by $O(\varepsilon)$ in magnitude (see equation (2.7)). We will use this requirement to have an idea of the thicknesses τ_1 and σ_2 of the boundary layers near the end-points $t = 0$ and $t = 1$ respectively. For this purpose we change the notations slightly and give the modified and unmodified solutions below.

Let $F = D(0,0)$ and U,V be the solutions of the unmodified and W, Z be the solutions of the modified boundary layer problems respectively.

Let $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$ $W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}$ $Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$

$U_1, W_1 \in \mathbb{R}^{n_1}$; $U_2, W_2 \in \mathbb{R}^{n_2}$;

$V_1, Z_1 \in \mathbb{R}^{n_3}$; $V_2, Z_2 \in \mathbb{R}^{n_4}$.

Then U, V, W, Z satisfy the following systems respectively.

$$dU/d\tau = FU \qquad 0 < \tau < \infty \qquad (2.10a)$$

$$U_1(0) = y_1(0) - Y_1(0) \qquad (2.10b)$$

$$\lim_{\tau \to \infty} U_2(\tau) = 0 \qquad (2.10c)$$

$$dV/d\sigma = -FV \qquad 0 < \sigma < \infty \qquad (2.11a)$$

$$V_2(0) = y_2(1) - Y_2(1) \qquad (2.11b)$$

$$\lim_{\sigma \to \infty} V_1(\sigma) = 0 \qquad (2.11c)$$

$$dW/d\tau = FW \qquad 0 < \tau < \tau_1 \qquad (2.12a)$$

$$W_1(0) = y_1(0) - Y_1(0) \qquad (2.12b)$$

$$W_2(\tau_1) = 0 \qquad (2.12c)$$

$$dZ/d\sigma = -FZ \qquad 0 < \sigma < \sigma_2 \qquad (2.13a)$$

$$Z_2(0) = y_2(1) - Y_2(1) \qquad (2.13b)$$

$$Z_1(\sigma_2) = 0 \qquad (2.13c)$$

In the following, we try to find out some sort of estimate of τ_1 by considering the systems (2.10) and (2.12). The results for σ_2 will follow similarly from the systems (2.11) and (2.13).

A crude estimate can be obtained by taking the solution U of (2.10) and requiring that $U_2(\varepsilon)$ passes through an ε-neighbourhood of $W_2(\tau)$ got from $W(\tau)$, the solution of (2.12). But, since (2.10) is a two-point boundary-value problem, we are till now not in a position to solve it explicitly. However, we might use some sort of shooting to achieve this. So assume that

(H-6) $U_2(0) = \alpha < \infty \quad , \quad \alpha \in \mathbb{R}^{n_4}$

such that $\lim_{\tau \to \infty} U_2(\tau) = 0$ holds.

Then, the solution of (2.10) is :

$$U(\tau) = e^{F\tau} \begin{bmatrix} y_1(0) - Y_1(0) \\ \alpha \end{bmatrix}$$

The requirement $|U(\tau_1)| \leq \varepsilon$, for some norm $|.|$, gives us the following.

Lemma 2.1

Under the assumption (H-6), τ_1 satisfies the inequality

$$\left| e^{F\tau_1} \begin{bmatrix} y_1(0) - Y_1(0) \\ \alpha \end{bmatrix} \right| \leq \varepsilon \qquad (2.14)$$

Moreover, if $e^{F\tau} = \begin{bmatrix} E_1(\tau) & E_2(\tau) \\ E_3(\tau) & E_4(\tau) \end{bmatrix}$ (2.15)

then $\alpha = -\lim_{\tau\to\infty} [E_4^{-1}(\tau)\, E_3(\tau)(y_1(0) - Y_1(0))]$ (2.16)

Now, looking from a different angle we find that the difference $\xi = U - W$ satisfies a system of the same type as (2.10). The only difficulty being the end-conditions. So assume that

(H-7) $W_2(0) = \beta < \infty$, $\beta \in \mathbb{R}^{n_4}$

such that $W_2(\tau_1) = O(\varepsilon)$ holds.

Then $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$, $\xi_1 \in \mathbb{R}^{n_3}$, $\xi_2 \in \mathbb{R}^{n_4}$

satisfies the following system :

$$d\xi/d\tau = F\xi \qquad 0 < \tau < \tau_1 \qquad (2.17a)$$

$$\xi_1(0) = 0,\ \xi_2(0) = \alpha - \beta \qquad (2.17b)$$

The solution of (2.17) is

$$\xi(\tau) = e^{F\tau} \begin{bmatrix} 0 \\ \alpha-\beta \end{bmatrix} .$$

Or, using (2.15) we can write it as

$$\xi(\tau) = \begin{bmatrix} E_2(\tau) & (\alpha-\beta) \\ E_4(\tau) & (\alpha-\beta) \end{bmatrix}$$

As before, we set $|\xi(\tau_1)| \le C_2 . \varepsilon$ to obtain

Lemma 2.2

Under the assumptions (H-6) - (H-7), τ_1 satisfies

$$\begin{vmatrix} E_2(\tau_1) & (\alpha-\beta) \\ E_4(\tau_1) & (\alpha-\beta) \end{vmatrix} \leq c_2.\varepsilon \qquad (2.18)$$

for some positive constant c_2, where α is determined by (2.16) and β satisfies :

$$| E_3(\tau_1)(y_1(0) - Y_1(0)) + E_4(\tau_1)\beta | \leq c_3.\varepsilon \qquad (2.19)$$

for some positive constant c_3.

It is of special significance to consider the degenerate cases arising from the trivialities $n_3 = 0$ and $n_4 = 0$. In these cases, the asymptotic solution of (2.1) - (2.2) satisfies a two-time scale property instead of the three-time-scales, that is because one of the boundary layers is lacking.

Case 1 When $n_3 = 0$.

Here, $y = y_2$. Consequently the boundary conditions (2.2) are given by $x_1(0)$, $x_2(1)$ and $y(1)$. The reduced system approximates the original one on $[0,1)$. At $t = 1$, the boundary-layer correction $V(\sigma)$ satisfies :

$$dV/d\sigma = -FV \qquad 0 < \sigma < \infty \qquad (2.20a)$$

$$V(0) = y(1) - Y(1) \qquad (2.20b)$$

$$\lim_{\sigma \to \infty} V(\sigma) = 0 \qquad (2.20c)$$

Setting $|V(\sigma_2)| \leq c_4 \cdot \varepsilon$, we obtain,

$$| e^{-F\sigma_2} (y(1) - Y(1)) | \leq c_4 \cdot \varepsilon \qquad (2.21)$$

for some positive constant $c_4 \cdot \varepsilon$

Case 2 When $n_4 = 0$.

In this case $y = y_1$. Thus we only consider the boundary layer correction at $t = 0$ to obtain the estimate (similar to Case 1) :

$$| e^{F\tau_1} (y(0) - Y(0)) | \leq c_5 \cdot \varepsilon \qquad (2.22)$$

for some positive constant c_5.

Remark 2.2

If both of n_3, n_4 are equal to zero, then (2.1) is not a singularly perturbed system and we might not observe the boundary layer phenomena. For non-homogeneous equations (unlike (2.1)), the same procedure might be applied, but it would involve the terms $e^{\int F(\eta) d\eta}$ which might need

extra hypotheses to yield practically meaningful results.

We finish this section by citing an example from O'Malley [83] where, the thickness is directly obtained.

Example 1

Consider the following TPBVP :

$$\begin{aligned} \dot{y}_1 &= y_2 \\ \varepsilon \dot{y}_2 &= -y_1 - y_2 \end{aligned} \qquad 0 < t < 1$$

with the boundary conditions,

$$y_1(0) = 0,\ y_2(1) = (\delta e^{\delta} - \mu e^{\mu})/(\delta\varepsilon - \mu\varepsilon)$$

where $\delta = (-1 + \sqrt{1-4\varepsilon})/2$,

$\mu = (-1 - \sqrt{1-4\varepsilon})/2$.

Writing the above system as a second order equation, we obtain, for $x = y_1$, $\dot{x} = y_2$;

$$\varepsilon \ddot{x} + \dot{x} + x = 0 .$$

The exact solution $x(t,\varepsilon)$ is given by :

$$x(t,\varepsilon) = (\delta\varepsilon - \mu\varepsilon)^{-1} (e^{\delta t} - e^{\mu t}).$$

If we neglect $O(\varepsilon)$ terms, we obtain,

$x(t,\varepsilon) \sim e^{-t} - e^{-t/\varepsilon}$, where $e^{-t/\varepsilon}$ represents the boundary-layer corrections. The thickness t_1 of the boundary layer near $t = 0$ can directly be obtained by taking $e^{-t_1/\varepsilon} \leq c_6$. for some constant c_6 at our disposal. Choosing $c_6 = 1$, we obtain $t_1 \geq |\varepsilon \ln \varepsilon|$.

4.3 Boundary Value Technique

In the introduction to this chapter (§ 4.1), we have remarked that some special numerical methods would be developed for systems of two-point boundary-value problems where there occurs no turning points. In the last section (§ 4.2) we found out the thickness of the possible boundary layers (in fact, the exponential of the thickness). This helps us to choose the points t_1, t_2 in $[0,1]$ properly so that the boundary layers might be expected to lie only within $[0,t_1]$ and $[t_2,1]$. In this section, we present such a method which will be discussed only for linear homogeneous TPBVP's. The non-homogeneous case can be handled similarly. To this end, we consider the following problem (see equation (2.1))

$$\begin{aligned} \dot{x} &= Ax + By \\ \varepsilon\dot{y} &= Cx + Dy \end{aligned} \qquad 0 < t < 1 \qquad (3.1a)$$

with boundary conditions

$$x_1(0) = a, \quad x_2(1) = b, \quad y_1(0) = c, \quad y_2(1) = d \qquad (3.1b)$$

where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$,

$x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, $y_1 \in \mathbb{R}^{n_3}$, $y_2 \in \mathbb{R}^{n_4}$;

$A(t,\varepsilon)$, $B(t,\varepsilon)$, $C(t,\varepsilon)$, $D(t,\varepsilon)$ are smooth matrix functions with appropriate dimensions and a, b, c, d are constant vectors, '.' denotes d/dt and ε is a small positive parameter. Though this class of problems is a very restricted one, this is worth paying an attention due to its appearance in many physical applications (e.g. see Chapter 6). We assume that there is a unique bounded solution to (3.1) and that there are no turning points or corner layers or interior layers in the solution to (3.1).

These assumptions lead us to a simplified grid-generation process in the hybrid method. In fact, if we could determine t_1 and t_2, the only possible stretching points,(i.e., where a finer grid is required),we need only some grid with a modest grid length so that the numerical methods we might

like to use are accurate. Note that we also need to determine the solution values at t_1 and t_2. In that case, we divide the interval of integration $[0,1]$ into three sub-intervals $[0,t_1]$, $[t_1,t_2]$ and $[t_2,1]$ and solve the stretched equations (according to the stretching parameters) on these sub-intervals independently. In the following, we describe this technique in detail and in a step-wise manner.

Step 1

Select points t_1 and t_2 in agreement with the inequalities of the form (2.18). Note that, usually the thickness of the boundary layers are $O(\varepsilon|\ln \varepsilon|)$. Determine the values of x and y at t_1 and t_2 using asymptotic expansion. To this end, let

$$x = \sum_{i=0}^{\infty} x_{(i)}(t)\ \varepsilon^i, \quad y = \sum_{i=0}^{\infty} y_{(i)}(t)\ \varepsilon^i \qquad (3.2)$$

where the subscript i within paranthesis shows the formal terms $x_{(i)}(t)$'s as t-dependent co-efficients of ε^i and etc. on substituting (3.2) in (3.1), the resulting equations can then be readily solved as initial value problems for $x_{(i)}(t)$ and $y_{(i)}(t)$; using analytical, or numerical means [58,28,112,82]. Once we get the values of $x(t_j)$, $y(t_j)$ for $j = 1,2$, either with $O(\varepsilon)$ or $O(\varepsilon^2)$ expansions, we switch

over to the next step, wherein we obtain a solution to the TPBVP (3.1).

Step 2

We scale the boundary layers by the factor ε. So denote

$$\tau(t) = t/\varepsilon \quad , \qquad \sigma(t) = (1-t)/\varepsilon \tag{3.3a}$$

$$\tau_1 = \tau(t_1), \qquad \sigma_2 = \sigma(t_2). \tag{3.3b}$$

Since, the boundary layers are guided by the τ- and σ-scaled equations, we resolve (3.1) into the following TPBVP's :

Left boundary layer :

$$dx/d\tau = \varepsilon(Ax + By) \tag{3.4a}$$

$$dy/d\tau = Cx + Dy \tag{3.4b}$$

on $0 < \tau < \tau_1$, with boundary conditions :

$$x_1(0) = a, \quad y_1(0) = c \tag{3.4c}$$

$$x_2(\tau_1) = x_2(t_1), \quad y_2(\tau_1) = y_2(t_1) \tag{3.4d}$$

Note that in (3.4d) above, the right hand side terms $x_2(t_1)$, $y_2(t_1)$ are as found out in Step 1.

Original system :

The equations (3.1a) with $t \in [t_1, t_2]$; the boundary conditions being

$x_1(t_1)$, $y_1(t_1)$ and $x_2(t_2)$, $y_2(t_2)$ as obtained in step 1.

Right boundary layer :

$$dx/d\sigma = -\varepsilon(Ax + By) \tag{3.5a}$$

$$dy/d\sigma = -Cx - Dy \tag{3.5b}$$

on $0 < \sigma < \sigma_2$, with boundary conditions :

$$x_2(0) = b , \qquad y_2(0) = d \tag{3.5c}$$

$$x_1(\sigma_2) = x_1(t_2) , \qquad y_1(\sigma_2) = y_1(t_2) \tag{3.5d}$$

Then the above three systems are solved numerically and are combined to give a uniformly valid solution on $[0,1]$. The left boundary layer and the right boundary layer are solved by using Euler's Implicit rule. We apply this rule, since the left and right boundary layers are stable in t and $-t$ respectively; the reverse stability of the right boundary layer is taken care of in formulating the system (3.5) by the transformation $\sigma = (1-t)/\varepsilon$. In the middle of the interval of integration, i.e.,

the original system as written above is solved by using trapezoidal rule. This agrees with the hybrid method, since on this interval the solution is slowly varying.

Remark 3.1

The boundary conditions in (3.4 - 3.5) and for the original system are chosen somewhat arbitrarily. This way of choosing the components x_1, x_2, etc. and recalling their values at t_1, t_2 comes in accord with the asymptotic choice (see the cancellation rules in O'Malley [82]). For example with (3.4a - 3.4b) we have to supply conditions at $t = 0$ and $t = \tau_1$. First, at $t = 0$, in (3.1b) we have the values $x_1(0)$ and $y_1(0)$. They are retained. Then we need the values of x_2 and y_2 at the points 0 or, τ_1. But they are obtained only at τ_1 in step 1. Hence the conditions (3.4c - 3.4d). Analogously boundary conditions in (3.5) are chosen. Then, the remaining boundary conditions are supplied to the original system.

Remark 3.2

Asymptotically, it is instructive to solve the reduced problem on the interval $[t_1, t_2]$. But this

creates further problem as to which boundary conditions are to be chosen; since the dimension of the reduced problem is n_1+n_2 compared to (3.1) which has the dimension $n_1+n_2+n_3+n_4$. The other reason is that, a better approximation is expected if we solve the original system instead of the reduced system.

Step 3

In choosing t_1 and t_2, note that we could only determine them upto an $O(\varepsilon|\ln\varepsilon|)$. So it is necessary to check (numerically) whether the choices are judicious. So let $X(t) = \begin{bmatrix} x \\ y \end{bmatrix}$; let $\dot{X}(t_1)_-$ and $\dot{X}(t_1)_+$ denote the approximations (as obtained from the numerical solution in step 2) of $\dot{X}(t)$ computed from equations (3.4) and the original system; let $\dot{X}(t_2)_-$ and $\dot{X}(t_2)_+$ denote that of $\dot{X}(t)$ at $t = t_2$ obtained from the original system and equations (3.5) respectively. Then use the following criteria. If,

$$|\dot{X}(t_i)_+ - \dot{X}(t_i)_-| \leq \delta, \qquad i = 1,2 \qquad (3.6)$$

is satisfied where $|\alpha|$ denotes the maximum of the absolute values of the components of α and δ is a prescribed tolerance limit (of $O(\varepsilon)$ or $O(\varepsilon^2)$), then we stop; else, we choose different t_1, t_2

still closer to $t = 0$ and $t = 1$ and proceed again from step 1 to step 3. In fact, we solve first by taking $t_1 = 10\,\varepsilon|\ln\varepsilon|$ and etc. and compare the norms in (3.6) for different iterations by making t_1 smaller (similarly for t_2), so that we can restrict ourselves to find better choices of t_1, t_2.

Remark 3.3

Since, the technique above heavily depends upon the choice of t_1 and t_2 ; and since the values of x and y at these points are determined using the asymptotic expansions in the boundary layers; we should not increase but decrease t_1 or $1-t_2$ in subsequent iterations. The name 'Boundary Value Technique' is given for this reason.

From among several test examples, the following is selected to illustrate the technique. The boundary conditions, are also specified for an even illustration.

Example 1

Consider the following second order problem (cf. [58])

$$\varepsilon\ddot{y} + (1+2t)\dot{y} + 2y = 0 \qquad 0 < t < 1 \quad (3.7a)$$

$$y(0) = 0, \quad y(1) = 1.$$

Writing this as a first order system, we obtain :

similarly the boundary conditions are supplied. The similar procedure is followed for $\varepsilon = 10^{-4}$, 10^{-5} and 10^{-6}.

The results for $\varepsilon = 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$ as obtained in the third iteration are summarised in Tables 4.1. In Table 4.1b, we give the comparison results, obtained by comparing the numerical solution produced by the boundary value technique with the asymptotic solution

$$y(t,\varepsilon) = 3\left(\frac{1}{1+2t} - e^{-\frac{t}{\varepsilon}}\right) + \varepsilon\left[\frac{2}{3}\left(\frac{1}{2t} + e^{-\frac{t}{\varepsilon}}\right) + 6\left(\frac{1}{(1+2t)^3} - e^{-\frac{t}{\varepsilon}}\right) - 3\frac{t^2}{\varepsilon^2}e^{-\frac{t}{\varepsilon}}\right] + O(\varepsilon^2) \quad (3.9)$$

In Figure 4.1 the solution $y(t)$ is shown when $\varepsilon = 10^{-5}$. Additional headings in the tables are as follows :

$$e_- = \max_{t \in BL} |u_A(t) - u_N(t)| \quad (3.10a)$$

$$e_o = \max_{t \in BB} |u_A(t) - u_N(t)| \quad (3.10b)$$

$$e_+ = \max_{t \in BR} |u_A(t) - u_N(t)| \quad (3.10c)$$

where BL, BR and BB are the left boundary layer, the right boundary layer and the region beyond the boundary layers respectively. $u_A(t)$ and $u_N(t)$ are the asymptotic and the numerical solutions respectively.

Table 4.1a

Choice of the Points t_1 and t_2 for Example - 1

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
t_1	4×10^{-3}	3.2×10^{-4}	4.4×10^{-5}	5×10^{-6}
t_2	no choice in 3rd iteration			

Table 4.1b

Summary of the Comparison Results (3rd iteration) for Example 1.

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e_-	0.2426E-1	0.6721E-2	0.1492E-3	0.3720E-4
e_o	0.3127E-2	0.7823E-3	0.2129E-3	0.4521E-4
e_+	No right boundary layer			

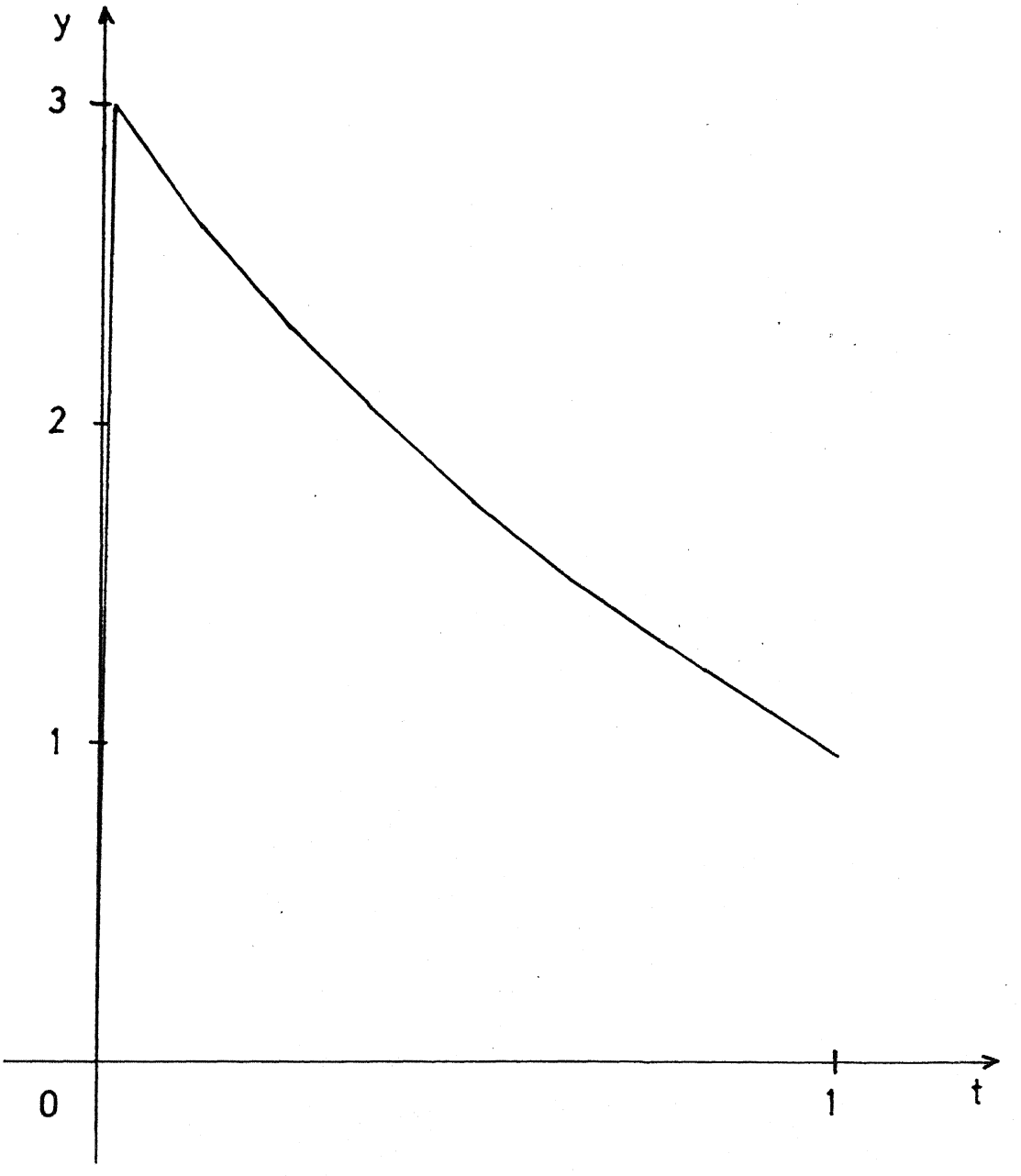


Figure 4.1

Numerical solution to example 1, $\varepsilon = 10^{-5}$

4.4 Cutting Point Technique

In § 4.3, we have presented the Boundary Value Technique to solve linear homogeneous singularly perturbed systems of two-point boundary value problems. The choice of the points t_1, t_2 that separate the inner regions and the outer regions (as they are usually called) is made partially by using the estimates in § 4.2 and modifying it appropriately by the successive iterations. Note that the estimates in § 4.2 yield an inequality of the form

$$c|\varepsilon \ln \varepsilon| \leq t_1 \qquad \text{for some constant } c.$$

First we choose greater values of t_1 and reduce it to such a value which would separate the inner regions and the outer region accurately.

In this section, we will follow the reverse trend, we first start from a value of t_1, (say, equal to ε) and increase it till we get the right value for the thickness of the boundary layer. Consequently during the computation the inner regions are made bigger and bigger in subsequent iterations so that the points t_1 and t_2 go inside the outer regions where asymptotically, the reduced solution is a 'good' approximation to the exact one.

For this reason, the first step in the Boundary Value Technique is replaced by the following which we write as step 1 (for the present method, the Cutting Point Technique).

Consider the equations (3.1). The Cutting Point Technique is described in a step-wise manner below.

Step 1

Choose $t_1, t_2 \in [0,1]$ close to 0 and 1 respectively i.e., in the neighbourhoods of the points 0 and 1 of radii of $O(\varepsilon|\ln \varepsilon|)$. For the evaluation of $x(t_i)$, $y(t_i)$, $i = 1,2$; take the reduced system corresponding to (3.1). The reduced equations are :

$$\dot{x} = Ax + By \qquad (4.1a)$$

$$0 = Cx + Dy \qquad (4.1b)$$

with the boundary conditions :

$$x_1(0) = a \ , \quad x_2(1) = b. \qquad (4.1c)$$

Here, we assume that (4.1) has a unique bounded solution. Then evaluate $x(t_i)$ and $y(t_i)$, $i = 1,2$ from (4.1) either by analytical or by numerical means.

Step 2

Step 2 of the Boundary Value Technique serves as step 2 of the Cutting Point Technique also.

Step 3

We require that (see step 3 of Boundary Value Technique in § 4.3)

$$| \dot{X}_+ (t_i) - \dot{X}_- (t_i) | < \delta , \quad i = 1, 2 \qquad (4.2)$$

If this requirement is not fulfilled, we choose t_1 and t_2 farther from 0 and 1, i.e., we increase the thicknesses t_1-0 and 1-t_2 and go to step 1 i.e., we choose t_1 greater than and t_2 less than that previously chosen and iterate step 1 to 3.

Among several test examples, the following example is chosen to demonstrate the method. It is noteworthy that, the method when applied to Example 1 in § 4.3 yielded an approximation in the 4th iteration which matches with the 3rd iteration of the Boundary Value Technique upto four significant decimal places.

Example 2

Consider the second order constant coefficient equation

$$\varepsilon \ddot{y} + \dot{y} + y = 0 \qquad 0 < t < 1 \qquad (4.3a)$$

$$y(0) = 0 , \quad \dot{y}(1) = 1/\varepsilon . \qquad (4.3b)$$

Or, in system form,

$$\dot{y} = z \qquad (4.4a)$$

$$\varepsilon \dot{z} = -y - z \qquad (4.4b)$$

on $0 < t < 1$ with boundary conditions,

$$y(0) = 0 , \quad z(1) = 1/\varepsilon . \qquad (4.4c)$$

The exact solution of (4.3) is :

$$y = c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t} \qquad (4.5)$$

with $\lambda_1 = (1 - \sqrt{1-4\varepsilon})/2\varepsilon$

$\lambda_2 = (1 + \sqrt{1-4\varepsilon})/2\varepsilon$

$c_1 = (\varepsilon c_3)^{-1}, \quad c_2 = -(\varepsilon c_3)^{-1}$

$c_3 = \lambda_2 e^{-\lambda_2} - \lambda_1 e^{-\lambda_1} .$

The results are summarised in Table 4.2 and the

solution is shown in Figure 4.2 for $\varepsilon = 10^{-4}$. Below, the tables are given for the 3rd iteration, where there is no need to choose t_2 as it appears from the first iteration that there is no right boundary layer.

Table 4.2a

Choice of the Point t_1 for Example 2

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
t_1	3×10^{-3}	4×10^{-4}	9×10^{-4}	5×10^{-6}

Table 4.2b

Summary of the Comparison Results (3rd iteration) for Example 2.

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e_-	0.4321E-2	0.2372E-3	0.4998E-4	0.1217E-4
e_o	0.5789E-2	0.2319E-2	0.1986E-3	0.4217E-4

Table 4.2a

Choice of the Point t_1 for Example 2

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
t_1	3×10^{-3}	4×10^{-4}	9×10^{-4}	5×10^{-6}

Table 4.2b

Summary of the Comparison Results (3rd iteration) for Example 2.

ε	10^{-3}	10^{-4}	10^{-5}	10^{-6}
e_-	0.4321E-2	0.2372E-3	0.4998E-4	0.1217E-4
e_o	0.5789E-2	0.2319E-2	0.1986E-3	0.4217E-4

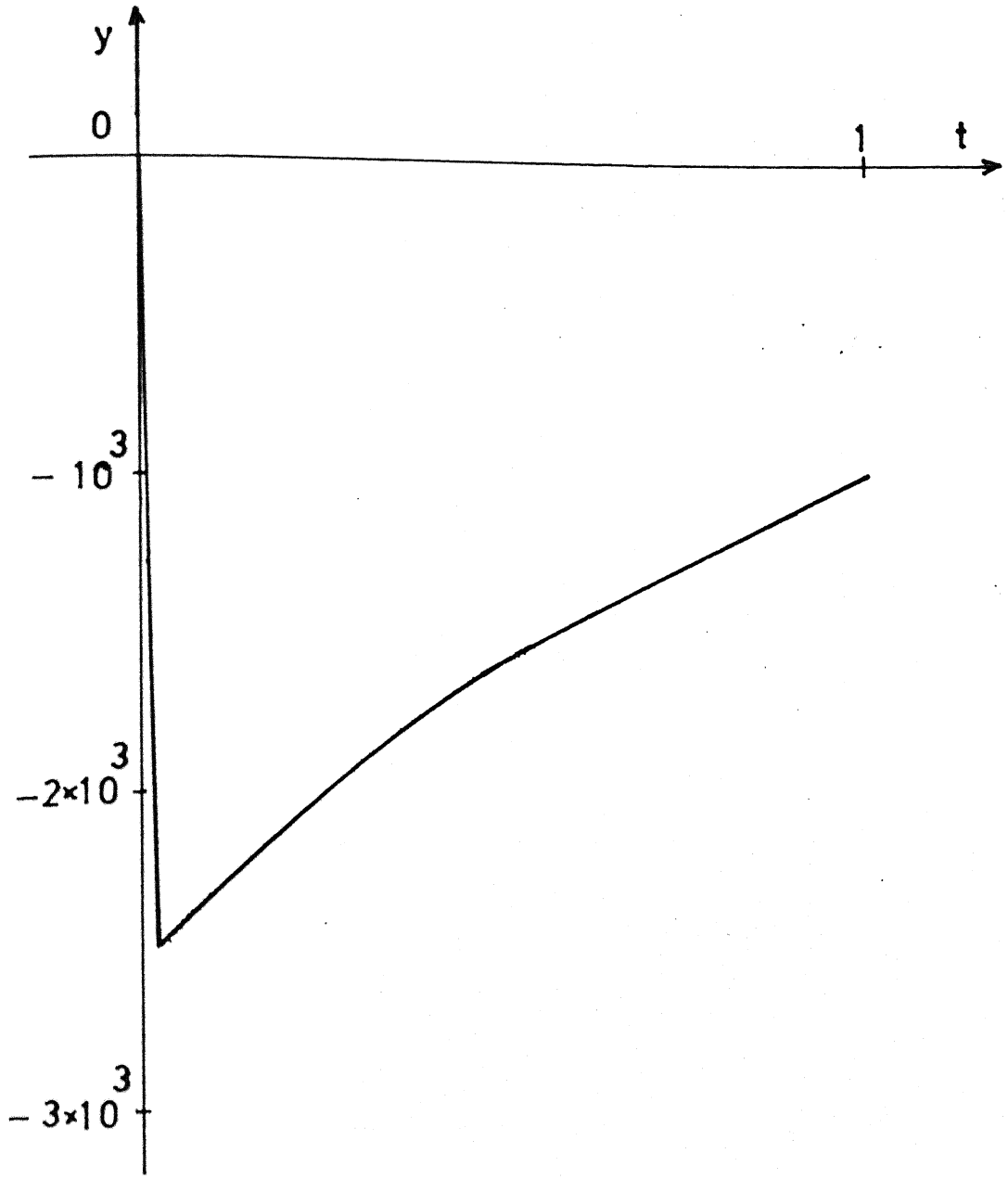


Figure 4.2

Numerical solution to example 2, $\varepsilon=10^{-4}$

4.5 Discussions

For the results of the numerical experimentations with the special methods, namely, the Boundary Value Technique and the Cutting Point Technique, we have seen that there is practically, a very little difference in the solutions obtained by the computational and asymptotic methods. This shows the reliability of these special methods. One should also look at the sub-class of problems these special methods tackle. These methods, we hope, might be applicable to the non-homogeneous linear systems but without interior layers and etc. We shall try latter to apply these methods for non-linear systems (see chapter 5). However, for non-homogeneous cases, the estimates in §4.2 are not sharp. So when we try to fix the size of the boundary layer regions (by trial and error) via the iterations, we must not go in one direction, i.e., we must not only decrease the boundary layer thicknesses throughout the computation nor that we must only increase it throughout the computation. That is, our choice of the thicknesses would oscillate from iteration to iteration. This way of trying for the thicknesses (determining appropriate t_1 and t_2) might be

incorporated in homogeneous cases also; which, ofcourse, will demand more machine time.

Another consideration in this direction is to use other methods, like Simpson's rule (e.g., Reddy and Kadalbajoo [51], Kellogg and Tsan [57]) in place of Euler's Implicit, and Trapezoidal rules. We have used these due to the analysis of the hybrid method in chapter 2.

Specifically for homogeneous systems a new technique would result, if we can eliminate some of the concerned variables to reduce the system into a second order system of lower dimension. This might be done in case the dimension of the system is (2m+2n). To be precise, consider the following system (which often arises in Optimal Control Theory; see Chapter 6).

$$\dot{x} = A_1x + A_2y + B_1p + B_2q \tag{5.1a}$$

$$\varepsilon\dot{y} = A_3x + A_4y + B_3p + B_4q \tag{5.1b}$$

$$\dot{p} = C_1x + C_2y + D_1p + D_2q \tag{5.1c}$$

$$\varepsilon\dot{q} = C_3x + C_4y + D_3p + D_4q \tag{5.1d}$$

on $0 < t < 1$ where $x,p \in \mathbb{R}^m$, $y,q \in \mathbb{R}^n$ and all the other matrices $A_i(t,\varepsilon)$, $B_i(t,\varepsilon)$, $C_i(t,\varepsilon)$ and

$D_i(t,\varepsilon)$ are of appropriate dimensions. For elegance, let

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \qquad B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

$$C = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \qquad D = \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}$$

and $$E = \begin{bmatrix} I_m & O \\ O & \varepsilon I_n \end{bmatrix}$$

I_j being the identity matrix of order j. Then, the system (5.1) might be rewritten by

$$E \frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + B \begin{bmatrix} p \\ q \end{bmatrix} \qquad (5.2a)$$

$$E \frac{d}{dt} \begin{bmatrix} p \\ q \end{bmatrix} = C \begin{bmatrix} x \\ y \end{bmatrix} + D \begin{bmatrix} p \\ q \end{bmatrix} \qquad (5.2b)$$

Assuming that B is non-singular throughout [0,1],

from (5.2a) we obtain

$$\begin{bmatrix} p \\ q \end{bmatrix} = B^{-1} \left[E \frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} - A \begin{bmatrix} x \\ y \end{bmatrix} \right] \qquad (5.3)$$

with $B^{-1} = \begin{bmatrix} F_1 & -F_2 \\ -F_3 & F_4 \end{bmatrix}$,

$$F_1 = (B_1 - B_2 B_4^{-1} B_3)^{-1}, \quad F_4 = (B_4 - B_3 B_1^{-1} B_2)^{-1},$$

$$F_2 = B_1^{-1} B_2 F_4, \qquad F_3 = B_4^{-1} B_3 F_1.$$

Putting (5.3) in (5.2b), we obtain the second order system

$$E \frac{d^2}{dt^2} \begin{bmatrix} x \\ y \end{bmatrix} = \dot{B} B^{-1} A \frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} + (\dot{A} - \dot{B} B^{-1} A) \begin{bmatrix} x \\ y \end{bmatrix} \qquad (5.4)$$

The same form could be obtained if instead we have D a non-singular matrix, where we can eliminate $\begin{bmatrix} x \\ y \end{bmatrix}$ instead of $\begin{bmatrix} p \\ q \end{bmatrix}$. The boundary conditions can be obtained for $\begin{bmatrix} x \\ y \end{bmatrix}$ by using (5.3).

For particular problems one might find which ever

way is convenient. The advantage in doing this is that one can apply the results for second order systems (see, e.g., Howes [45], Eckhaus [28] and Engstrom [29]). Also, the interval of integration can be taken as any bounded closed interval $[t_o, t_f]$ in place of $[0,1]$.

In summary, we are justified in trying to find a solution in the left boundary layer with the stretching transformation $\tau = t/\varepsilon$ and in the right boundary layer, with $\sigma = (1-t)/\varepsilon$, since the boundary layers are stable in time t and reverse time $-t$ respectively. Again, the approximation in the outer regions (i.e., beyond the boundary layers) by the proposed methods are expected to be better than the asymptotic solutions, since in the proposed methods, the original system is solved, whereas in the asymptotic techniques, the solution in the outer region is represented totally by the reduced solution. One is justified to approximate the solution in this outer region by the reduced solution, when ε is small enough, i.e., an infinitesimal (see e.g., [73] and etc.; also a beautiful counter example in Kokotovic and Yackel [61], when $\varepsilon = 0.1$). But in practice, ε is not an infinitesimal; whence it

amounts to make the proposed methods preferable. Also, since the thickness of the boundary layers is not yet determined with analytical or asymptotic means for the systems in hand, and asymptotic methods available to solve this class of problems almost avoid this question, it is suggestive that we try to get a solution by these special methods, where we would be able to know more accurately about the thickness of the boundary layers (rather than just $O(\varepsilon)$ or $O(\varepsilon |\ln \varepsilon|)$). To note a few, among the merits, the methods do not require much skill necessary in the conventional methods (e.g., matching the coefficients in the asymptotic expansions), it does not require very fine mesh size; the methods require very less problem preparation for computer adaptation , these are also efficient and accurate; the numerical experiments demonstrate this fact. It might also be observed that an accuracy predicted could always be achieved with comparatively less computational effort.

Chapter 5

NONLINEAR SINGULARLY PERTURBED TWO-POINT BOUNDARY-VALUE SYSTEMS

5.1 Introduction

So far we have presented some efficient and accurate methods for linear systems of singularly perturbed two-point boundary-value problems. In this chapter, we consider the corresponding nonlinear problems of the form :

$$\dot{x} = f(x,y,t,\varepsilon) \qquad (1.1a)$$

$$0 < t < 1$$

$$\varepsilon\dot{y} = g(x,y,t,\varepsilon) \qquad (1.1b)$$

with boundary conditions

$$B(x(0), x(1), y(0), y(1),\varepsilon) = 0 , \qquad (1.1c)$$

$x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, f, g are smooth nonlinear functions; (1.1c) represents the boundary condition, which when linearized would provide m+n linearly independent conditions. We also assume the uniqueness and boundedness of the solution to (1.1) on the interval [0,1]. For numerical solution of (1.1) one often thinks of applying usual methods ([55-56]). The experience with the linear case shows

that the usual two-point boundary-value methods are not expected to work satisfactorily. Whatever the methods are, one should try to find out estimates of the type discussed in Chapter 2. But this certainly aims at too much of a difficult task, as the linearity, the most important of all, is lacking.

This leads us to consider the other approach, i.e., using iterations. The procedure is that we linearize the problem (1.1) first and then find out numerical solution by the methods developed in the preceding chapters. Though this is quite plausible, the main disadvantage of this approach is in terms of cost. It approximates the system (1.1) by a sequence of linear systems; all of which then must be solved numerically. However, experience with some specific problems of the class defined by (1.1) shows that we might follow this approach.

In § 5.2 such an iteration process, i.e., a modification of the familiar Newton's iteration is proposed. Then a model example is considered in § 5.3, and the method thus developed is applied to this example. And in the last section, the remarks on the efficiency of the methods are made.

5.2 Modified Newton's Iteration

For notational convenience, we consider the nonlinear autonomous equations corresponding to Equations (1.1); we also supress the explicit appearance of ε in (1.1), for the time being. So we consider the system

$$\dot{z} = F(z) \qquad 0 < t < 1 \qquad (2.1a)$$

with non-linear boundary condition whose linearization yields

$$B_o z(0) + B_1 z(1) = B_2 \qquad (2.1b)$$

representing k linearly independent boundary conditions; k the dimension of z.

The Newton's iteration is defined in Theorem (6.6) of Chapter 1 where also the quadratic convergence is stated. However, for practical purposes, the Frechet derivatives or the G-derivatives do not play very important role. So a modification of the iteration can be carried out by replacing the Frechet or G-derivatives by the corresponding Jacobian ∂F. To this end, we use the generalized Taylor's theorem for operators (Theorem 6.3 of Chapter 1) to obtain

$$||P(z) - P(z_o) - \partial P(z_o)(z-z_o)||$$

$$\leq \sup_{w \in L(z_o,z)} \frac{1}{2} ||\partial^2 P(w)||.||z-z_o||^2, \quad (2.2)$$

by taking $n = 2$ in Theorem (6.3) of Chapter 1. Note that $\partial P(z_o)$ and $\partial^2 P(w)$ denote $\partial P(z_o)/\partial z_o$ and $\partial^2 P(w)/\partial w^2$ respectively. Newton's iteration might be written (for this case) by :

$$z_{n+1} = z_n + \bar{z}_n \qquad n = 0, 1, 2, ... \quad (2.3a)$$

where $\bar{z}_n$ satisfies :

$$\partial P(z_n)\, \bar{z}_n = - P(z_n) \; ; \quad (2.3b)$$

The subscripts n in (2.3) denote the n th approximation of z, the solution of the equation

$$P(z) = 0 \quad (2.4)$$

It follows that the iteration (2.3) converges quadratically near z; since, (2.2 - 2.4) implies that

$$||z_{n+1} - z|| \leq ||[\partial P(z_n)]^{-1}|| \cdot \sup_{w \in L(z_n,z)} \frac{1}{2} ||\partial^2 P(w)||.||z_n-z||^2 ,$$

assuming that $P(w)$, $[\partial P(w)]^{-1}$ and $\partial^2 P(w)$ are

bounded in some neighbourhood $N(z,r)$ with $P(z) = 0$. Now, our aim is to apply the iteration (2.3) to (2.1).

Here, instead of applying the iterations above directly by letting $P(w) \equiv \dot{w} - F(w)$, we will go a little bit further. Suppose that the nth approximation z_n is already determined. We want to find out the (n+1)th approximation z_{n+1}. But z_{n+1} satisfies (2.1a) which means :

$$\dot{z}_{n+1} = F(z_{n+1}) . \tag{2.5}$$

Our scheme is to approximate $F(z_{n+1})$ by the method of operators, i.e., we let

$$F(z_{n+1}) \sim F(z_n) + \partial F(z_n)(z_{n+1} - z_n). \tag{2.6}$$

Note that, formally we are not justified in approximating this way. Assuming that this linearization is acceptable, we obtain :

$$\dot{z}_{n+1} = F(z_n) + \partial F(z_n)(z_{n+1} - z_n) \tag{2.7}$$

Subtracting $\dot{z}_n$ from each side of (2.7), we get

$$\dot{z}_{n+1} - \dot{z}_n = \partial F(z_n)(z_{n+1} - z_n) + (F(z_n) - \dot{z}_n) \tag{2.8}$$

Denoting $\bar{z}_n = z_{n+1} - z_n$, we obtain

$$\dot{\bar{z}}_n = \partial F(z_n)\, \bar{z}_n + (F(z_n) - \dot{z}_n) \tag{2.9}$$

Now, (2.3a) and (2.9) define the iteration process, we are interested in .

Remark 2.1

We have avoided the second derivatives that might have come up in considering the operator $P(w) = \dot{w} - F(w)$ directly. The second point is that, one should not neglect the term $F(z_n) - \dot{z}_n$ in (2.9), no matter how much temptation it might offer, because, z_n might not be approximating the exact solution.

In the following we shall show that the convergence of the modified Newton's iteration (2.3a) and (2.9) is also quadratic. We begin with an initial value problem (2.1a - 2.1b) and

$$z(0) = \alpha < \infty . \qquad (2.10)$$

Very often in Mathematics, it happens that the same result is proved again and again. This seems futile but it is not. Our proof of convergence below begs the same excuse.

Lemma 2.1

Let $w(t) \in \mathbb{R}^k$ be the solution of the initial value problem

$$\dot{w} = \partial F(y)(w-y) + F(y) \qquad 0 < t \leq 1 \qquad (2.11a)$$

$$w(0) = y(0) \qquad (2.11b)$$

$y(t)$ being a known function of t. Assume that

(a) $\sup_{x \in L(y,w)} ||\partial^2 F(x)|| = O(1)$

(b) $\exp\left(\int_{\eta}^{t} ||\partial F(y(\xi))|| d\xi \right) = O(1)$, $0 \leq \eta \leq t \leq 1$

Then, $||F(w) - \dot{w}|| \leq K_1 = O\left(\sup_{0 \leq t \leq 1} ||F(y) - \dot{y}||^2\right)$.

Proof (2.11) implies ([94], Sec.33) that

$$||w(t) - y(t)||$$

$$\leq \int_{0}^{t} \exp\left(\int_{\eta}^{t} ||\partial F(y(\xi))|| d\xi \right) ||F(y(\eta)) - \dot{y}(\eta)|| d\eta$$

$$\leq \sup_{0 \leq t \leq 1} ||F(y) - \dot{y}|| \int_{0}^{t} \exp\left(\int_{\eta}^{t} ||\partial F(y(\xi))|| d\xi \right) d\eta$$

$$\leq \sup_{0 \leq t \leq 1} ||F(y) - \dot{y}|| \sup_{0 \leq \eta \leq t} \left(t \exp\left(\int_{\eta}^{t} ||\partial F(y(\xi))|| d\xi \right) \right)$$

i.e.,

$$||w(t) - y(t)|| \leq K_2 = O\left(\sup_{0 \leq t \leq 1} ||F(y) - \dot{y}||\right).$$

Again,

$$||F(w) - \dot{w}||$$

$$= ||f(w) - f(y) - \partial F(y)(w-y)||$$

$$\leq \frac{1}{2} \times \sup_{\xi \in L(y,w)} ||\partial^2 F(x)|| \; ||w-y||^2$$

from which the result follows.

Lemma 2.2

If $z(t) \in \mathbb{R}^k$ is a solution of

$$\dot{z} = F(z) \qquad 0 < t \leq 1, \quad z(0) \text{ given} ; \tag{2.12}$$

and $w(t) \in \mathbb{R}^k$ is a solution of

$$\dot{w} = \partial F(y)(w-y) + F(y) \qquad 0 < t \leq 1 \tag{2.13a}$$

$$w(0) = z(0) \tag{2.13b}$$

with a given function y in a neighbourhood $N(z,r)$, $r > 0$ of z and conditions (a) - (b) in Lemma 2.1 are satisfied, then

$$||z-w|| \leq K = O(||z-y||^2).$$

Proof

Let $s = z-w$

and $G(y,z) = F(z) - F(y) - \partial F(y)(z-y)$

then from Theorem 6.3 of Chapter 1, we obtain

$$||G(y,z)|| \leq \frac{1}{2} \sup_{x \in L(y,z)} ||\partial^2 F(x)||.||z-y||^2.$$

Also, from (2.12) - (2.13) we obtain

$$\dot{s} = \partial F(y)s + G(y,z), \quad s(0) = 0.$$

Then, as in Lemma 2.1

$$||s(t)|| \leq \int_0^t \exp\left(\int_\eta^t ||\partial F(y(\xi))|| d\xi \right) . ||G(y(\eta))\ z(\eta)|| d\eta$$

$$\leq \frac{1}{2} \sup_{x \in L(y,z)} ||\partial^2 F(x)|| t \exp\left(\int_\eta^t ||\partial F(y(\xi))|| d\xi\right) . \sup_{0 \leq \eta \leq t} ||z(\eta) - y(\eta)||^2$$

which proves the lemma.

Immediately, we get the following theorem which states that the solution of (2.1) with an initial value $z(0)$ obtained by iteration has quadratic convergence.

Theorem 2.1

Assume that the conditions (a) - (b) in Lemma 2.1 are satisfied, then the solution $z^* (= \lim_{n \to \infty} z_n)$ obtained by the following iteration scheme :

$$\bar{z}_n = \partial F(z_n)\ \bar{z}_n + (F(z_n) - \dot{z}_n) \qquad (2.14a)$$

$$\bar{\dot{z}}_n = 0 \qquad (2.14b)$$

$$z_{n+1} = z_n + \bar{z}_n \qquad (2.14c)$$

z_o suitably chosen with $z_o(0) = z(0)$ (2.14d)

converges quadratically to the solution $z(t)$ of the equations

$$\dot{z}(t) = F(z) \qquad 0 < t \leq 1 \qquad (2.15a)$$

$z(0)$ given. (2.15b)

Moreover, if $||F(z_o) - \dot{z}_o|| < 1$, then (2.14) solves (2.15) with $||F(z_n) - \dot{z}_n||$ converging quadratically to zero.

Proof

In Lemma 2.2 put $w = z_{n+1}$, $y = z_n$ to get the algorithm (2.14). We also obtain

$||z - z_{n+1}|| \leq K = O(||z - z_n||^2)$, proving the first part of the theorem. By taking $w = z_{n+1}$, $y = z_n$ in Lemma 2.1, we obtain

$$||F(z_{n+1}) - \dot{z}_n| \leq K_1 = O(||F(z_n) - \dot{z}_n||^2)$$

which proves the second part.

Instead of the given initial value $z(0)$ suppose that a separated set of boundary conditions

is given, e.g.,

$$z_1(0) = \beta, \quad z_2(1) = \gamma, \quad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{2.16}$$

Then, also the lemmas hold, since in that case, we replace the integrals of the type $\int_0^t \varphi(\ldots)$ by

$$\begin{bmatrix} \int_0^t \varphi_1(\ldots) \\ \int_1^t \varphi_2(\ldots) \end{bmatrix} \quad \text{where} \quad \varphi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} .$$

However, for the TPBVP (2.1) we assume that it possesses a unique bounded solution. Then, since $z(t)$ is bounded, we have, in particular, $z(0) = \alpha < \infty$. Then the Theorem 2.1 will hold. We note this down in the following..

Theorem 2.2

If the solution of (2.1) is unique and bounded and F satisfies the conditions (a) - (b) in Lemma 2.1, then the iteration scheme (2.14a - 2.14c) with z_o satisfying (2.16) solves (2.1a-c) with quadratic convergence.

Remark 2.2

The condition $\bar{z}_n = 0$ in (2.14) might be

replaced by $B_o \bar{z}_n(0) + B_1 \bar{z}_n(1) = C$ where $B_o z(0) + B_1 z(1) = B_2$ is to be satisfied by the initial approximation z_o. This way of replacement should be carried out to solve (2.1) numerically.

Note that we have considered only autonomous systems. However, it is standard to extend the theorems also to non-autonomous equations. Suppose that instead of (2.1) the following equation is considered

$$\dot{z} = F(t,z). \tag{2.17}$$

Let $w = \begin{bmatrix} t \\ z \end{bmatrix}$, $G(w) = \begin{bmatrix} 1 \\ F(t,z) \end{bmatrix}$.

Then equation (2.17) is transformed to

$$\dot{w} = G(w),$$

which is autonomous. Similarly we add $t(0) = 0$ with the boundary conditions and then apply Theorem 2.2.

To solve the linear systems in the iteration steps (in (2.14)) we apply the hybrid method developed in Chapter 2. The other methods, i.e., the special methods of Chapter 4 might also be applied; but in that case, one must see that there are no interior layers in the original non-linear problem.

5.3 A Model Nonlinear Example

In the previous section, we have discussed a modification of the Newton's iteration we intend to apply for the nonlinear singularly perturbed two-point boundary-value problems. The systems we are concerned with are of the form :

$$\dot{x} = f(x, y, t, \varepsilon) \tag{3.1a}$$

$$\varepsilon\dot{y} = g(x,y,t, \varepsilon) \tag{3.1b}$$

on $0 < t < 1$ with boundary conditions

$$B(x(0), x(1), y(0), y(1), \varepsilon) = 0 \tag{3.1c}$$

with the usual conventions.

$$\text{Writing } z = \begin{bmatrix} x \\ t \\ y \end{bmatrix}, \; G = \begin{bmatrix} f \\ 1 \\ \varepsilon^{-1}g \end{bmatrix} \tag{3.2}$$

(3.1a - 3.1b) can be reduced to

$$\dot{z} = G(z) \tag{3.3}$$

Our aim is to write the linearization directly for (3.1). For this purpose, we first compute ∂G.

$$\partial G = \begin{bmatrix} \partial\bar{f}/\partial w & \partial\bar{f}/\partial y \\ \varepsilon^{-1}\partial g/\partial w & \varepsilon^{-1}\partial g/\partial y \end{bmatrix}$$

where $w = \begin{bmatrix} x \\ t \end{bmatrix}$ $f = \begin{bmatrix} \bar{f} \\ 1 \end{bmatrix}$.

Let $\bar{w}_n = w_{n+1} - w_n$ and $\bar{y}_n = y_{n+1} - y_n$. Then the linearization (2.14) reduces to :

For $n = 0,1,2,\ldots$

$$\dot{\bar{w}}_n = \frac{\partial \bar{f}}{\partial w}(w_n,y_n)\bar{w}_n + \frac{\partial \bar{f}}{\partial y}(w_n,y_n)\bar{y}_n + (f(w_n,y_n) - \dot{w}_n) \tag{3.4a}$$

$$\varepsilon \dot{\bar{y}}_n = \frac{\partial g}{\partial w}(w_n,y_n)\bar{w}_n + \frac{\partial g}{\partial y}(w_n,y_n)\bar{y}_n + (g(w_n,y_n) - \dot{y}_n) \tag{3.4b}$$

$$w_{n+1} = w_n + \bar{w}_n \tag{3.4c}$$

$$y_{n+1} = y_n + \bar{y}_n \tag{3.4d}$$

w_o and y_o are initialized so as to satisfy a linearization of (3.1c), i.e.,

$$B_o \begin{bmatrix} x_o(0) \\ y_o(0) \end{bmatrix} + B_1 \begin{bmatrix} x_o(1) \\ y_o(1) \end{bmatrix} = B_2 \tag{3.4e}$$

with $t_o(0) = 0.$ (3.4f)

Note that here we might also choose the initial functions so that

$||F(z_o) - \dot{z}_o|| < 1$ to obtain the results of Lemma 2.2. Now, (3.4a - 3.4b) and (3.4e - 3.4f) are the linear systems we handled in Chapter 3. So we apply the hybrid method to solve the systems (3.4a - 3.4b) and (3.4e - 3.4f) for each n. As experience shows it only takes 3 to 4 iterations to obtain an accurate numerical solution. However, the choice of the initial approximation z_o is significant in reducing the number of iterations for achieving certain predicted accuracy.

The purpose of this section is to apply this numerical technique to a model non-linear example taken from Kevorkian and Cole [58], where almost all the varied nature of the solution are present.

For computational point of view, we add another step to the above iteration and solution by the hybrid method. Note that, it is not yet clear, where to stop the iteration. To this end, we require an inequality of the type

$$||z_{n+1} - z_n|| < \delta \qquad (3.5)$$

to hold, where δ is a tolerence limit, which we prefer to take $O(h^2)$, where h is the maximum grid length (matching with the quadratic convergence

of the Newton's iteration and with the accuracy of the hybrid method).

The example we want to study is :

$$\varepsilon \ddot{x} + x\dot{x} - x = 0 \qquad 0 < t < 1 \qquad (3.6a)$$

$$x(0) = A, \quad x(1) = B \; . \qquad (3.6b)$$

$x(t) \in \mathbb{R}$; A and B will be specified later. Equation (3.6a) is invariant under the transformation :

$$x \longrightarrow -x, \qquad t \longrightarrow 1-t, \qquad A \longrightarrow -B \; .$$

Thus, if $x = f(t,A,B)$ is a solution of (3.6), then $x = -f(1-t, -B, -A)$ is also a solution and the later solution satisfies the boundary conditions :

$$x(0) = -B, \qquad x(1) = -A \; .$$

A solution for a given point (A,B) generates a solution for the reflected point (-B, -A). This reflected solution is obtained by the transformation

$$x \longrightarrow -x, \qquad t \longrightarrow 1-t$$

i.e., a reflection about the t-axis followed by a reflection about '$t = \frac{1}{2}$' axis. Hence we need only consider half of the A,B-plane, i.e., $B \geq -A$.

In the region $B < -A$, the reflected solution will hold. We have the following distinct cases displaying varied nature of the solution. (Kevorkian and Cole [58]). The cases are :

Case -- 0. $B = A + 1$

Case - 1. $A > B-1 > 0$

Case -- 2. $0 \leq |A| < B-1$

Case - 3. $B > A+1, \quad -(B+1) < A < 1-B.$

Case -- 4. $B-1 < A < 0, \quad 0 < B < A+1$

Case - 5. $0 < B < 1, \quad A > 0.$

Case - 6. $B < 0 < A$

First, we cast out the trivial case $A = B = 0$; where $x = 0$ is the solution to (3.6). The two outer solutions satisfying the two boundary conditions respectively are

$$x_R(t) = t + B - 1 \tag{3.7}$$

$$x_L(t) = t + A \tag{3.8}$$

Case - 0. $B = A + 1.$

The two outer solutions are the same, and they represent the exact solution.

Case - 1. $A > B-1 > 0$.

The outer solution $x_R = t + B - 1$ satisfies the right boundary condition and the boundary layer matched to x_R is :

$$B_L(\tau) = (B-1) \coth \left(\frac{B-1}{2}(\tau + C)\right)$$

$$\tau = t/\varepsilon \ , \quad C = \frac{2}{B-1} \coth^{-1} \left(\frac{A}{B-1}\right).$$

Case - 2. $0 \leq |A| < B-1$.

The outer solution x_R satisfies the right boundary condition and the left boundary layer is given by :

$$B_L(\tau) = (B-1) \tanh \left[\frac{B-1}{2} (\tau + C)\right] ,$$

$$C = \frac{2}{B-1} \tanh^{-1} \left(\frac{A}{B-1}\right) .$$

Case - 3. $B > A+1, \quad -(B+1) < A < 1-B$.

Let $\bar{t} = (1 - A - B)/2$.

The outer solutions x_L and x_R satisfy both the boundary conditions respectively. The inner solution at $\bar{t}$ is :

$$I(\eta) = \frac{B-A-1}{2} \tanh \left(\frac{B-A-1}{4} \eta\right) ,$$

$$\eta = (t - \bar{t})/\varepsilon \ . \qquad \text{(a shock layer)}$$

Case - 4. $B - 1 < A < 0,\ 0 < B < A + 1.$

The outer solution has three pieces :

$$x_L(t) = t+A \qquad 0 \le t \le -A$$

$$x_M(t) = 0 \qquad -A < t < 1-B$$

$$x_R(t) = t-1+B \qquad 1-B \le t \le 1$$

Here f_{LC} provides the match between x_L, x_M ; and f_{RC}, the match between x_M, x_R ; where f_{LC} and f_{RC} are two exponential functions satisfying

$$\lim_{\tau\to\infty} f_{LC}(\tau) = 0 , \quad \lim_{\tau\to-\infty} f_{LC}(\tau) = \tau$$

$$\lim_{\tau\to\infty} f_{RC}(\tau) = \tau , \quad \lim_{\tau\to-\infty} f_{RC}(\tau) = 0 .$$

Note that f_{LC} and f_{RC} only smoothen the corners present in the solution.

Case - 5. $0 < B < 1, \qquad A > 0$

Write $f_{LT} = \dfrac{2A\varepsilon}{A^t + 2\varepsilon}$, and f_{RT} , the reflection of f_{LT}. In this case, f_{LT} brings the solution from A to $x = 0$; f_{RC} brings the solution from $x = 0$ to x_R.

Case - 6.

Here, the outer solution is $x_M = 0$; which is joined at the left by f_{LT} and to the right by f_{RT} to $A > 0$ and $B < 0$ respectively.

Note that the equation (3.6a) can be written in system form as :

$$\dot{x} = y \tag{3.7a}$$

$$\varepsilon\dot{y} = x - xy \tag{3.7b}$$

And, then the corresponding linearization is :

$$\dot{w}_n = v_n \tag{3.8a}$$

$$\varepsilon\dot{v}_n = -x_n v_n + (1-y_n)w_n + (x_n - x_n y_n) \tag{3.8b}$$

$$w_n(0) = w_n(1) = 0 \tag{3.8c}$$

with $$x_{n+1} = x_n + w_n \tag{3.8d}$$

$$y_{n+1} = y_n + v_n \ , \tag{3.8e}$$

where we assume that we have already initialized the functions x_o and y_o. For example, in Case 1, we let $x_o(0) = A$, $x_o(1) = B$; which of-course could be taken for all the cases; however, the numerical experimentation shows that there are certain advantages in initializing other functions instead of the obvious one, i.e., a straight line connecting $x_o(0) = A$ and $x_o(1) = B$.

In the following, we only present the results for Case-2, Case-3 and Case-4. Case-0 is omitted

Mid : Minimum of the absolute values of the differences between the asymptotic solution and the numerical solution.

Nit : Specifies the iteration number, i.e., 'n' in 'nth iteration'.

Table 5.1a

Choice of Stretching Points (Case-2)

	ε	1E-2	1E-4	1E-6
	Nit	4	3	3
SP:	b_1	0.036	0.001	2E-5
	b_2	0.96	0.996	99994E-5

Table 5.1b

Summary of the Comparison Results (Case-2)

ε	SS	$[0,b_1]$	$[b_1,b_2]$	$[b_2,1]$
1E-2	NG	8	16	8
	Mad	24E-3	88E-3	28E-4
	Mid	OE-1	2E-6	OE-1
1E-4	NG	16	16	8
	Mad	29E-5	32E-5	4E-4
	Mid	OE-1	3E-6	OE-1
1E-6	NG	12	20	8
	Mad	31E-6	15E-6	1E-5
	Mid	OE-1	5E-8	OE-1

Table 5.2a

Choice of Stretching Points (Case-3)

	ε	1E-2	1E-4	1E-6
	Nit	3	4	5
SP :	b_1	0.46	0.499	0.49993
	b_2	0.522	0.504	0.50018

Table 5.2b

Summary of Comparison Results (Case-3)

ε	SS	$[0,b_1]$	$[b_1,b_2]$	$[b_2,1]$
1E-2	NG	12	16	12
	Mad	2E-2	6E-3	21E-3
	Mid	OE-1	2E-6	OE-1
1E-4	NG	12	16	12
	Mad	2E-3	18E-4	4E-4
	Mid	OE-1	25E-7	OE-1
1E-6	NG	16	16	20
	Mad	2E-5	22E-6	3E-5
	Mid	OE-1	6E-8	OE-1

Table 5.3a

Choice of Stretching Points (Case-4)

	ε	1E-2	1E-4	1E-6
	Nit	4	4	4
SP :	b_1	0.498	0.4999	0.49987
	b_2	0.501	0.5002	0.50002
	b_3	0.665	0.6665	0.66661
	b_4	0.668	0.6667	0.66667

Table 5.3b

Summary of Comparison Results (Case-4)

ε	SS	$[0,b_1]$	$[b_1,b_2]$	$[b_2,b_3]$	$[b_3,b_4]$	$[b_4,1]$
	NG	8	8	8	8	8
1E-2	Mad	2E-4	2E-2	2E-6	4E-3	4E-6
	Mid	OE-1	2E-7	OE-1	3E-7	OE-1
	NG	8	12	8	12	8
1E-4	Mad	12E-6	16E-4	11E-5	2E-4	4E-7
	Mid	OE-1	25E-8	OE-1	27E-8	OE-1
	NG	8	16	8	16	8
1E-6	Mad	2E-7	54E-8	3E-8	22E-6	14E-8
	Mid	OE-1	12E-8	OE-1	9E-8	OE-1

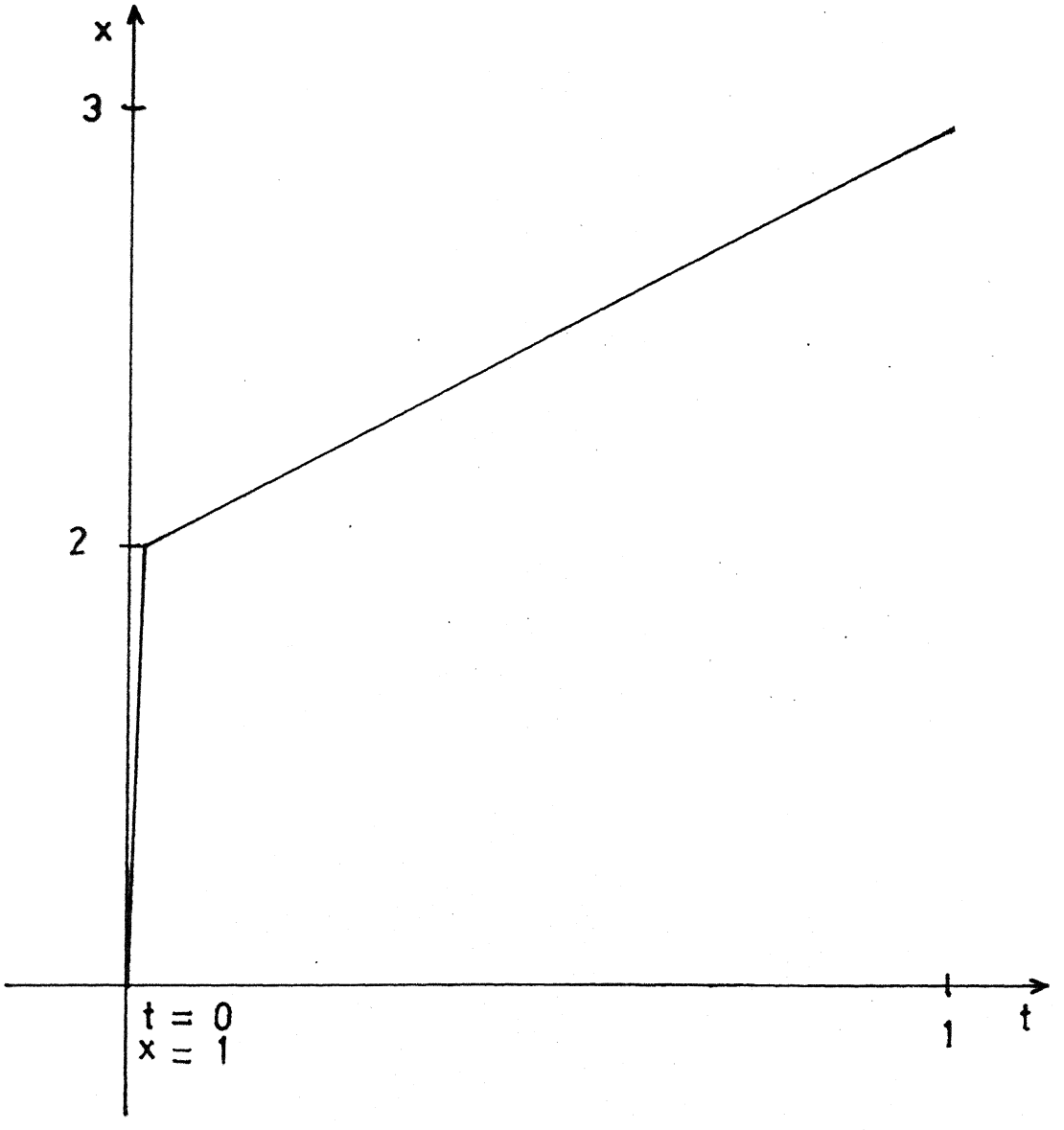


Figure 5 . 1

Numerical solution to case-2

$\varepsilon = 10^{-4}$, 3rd iteration.

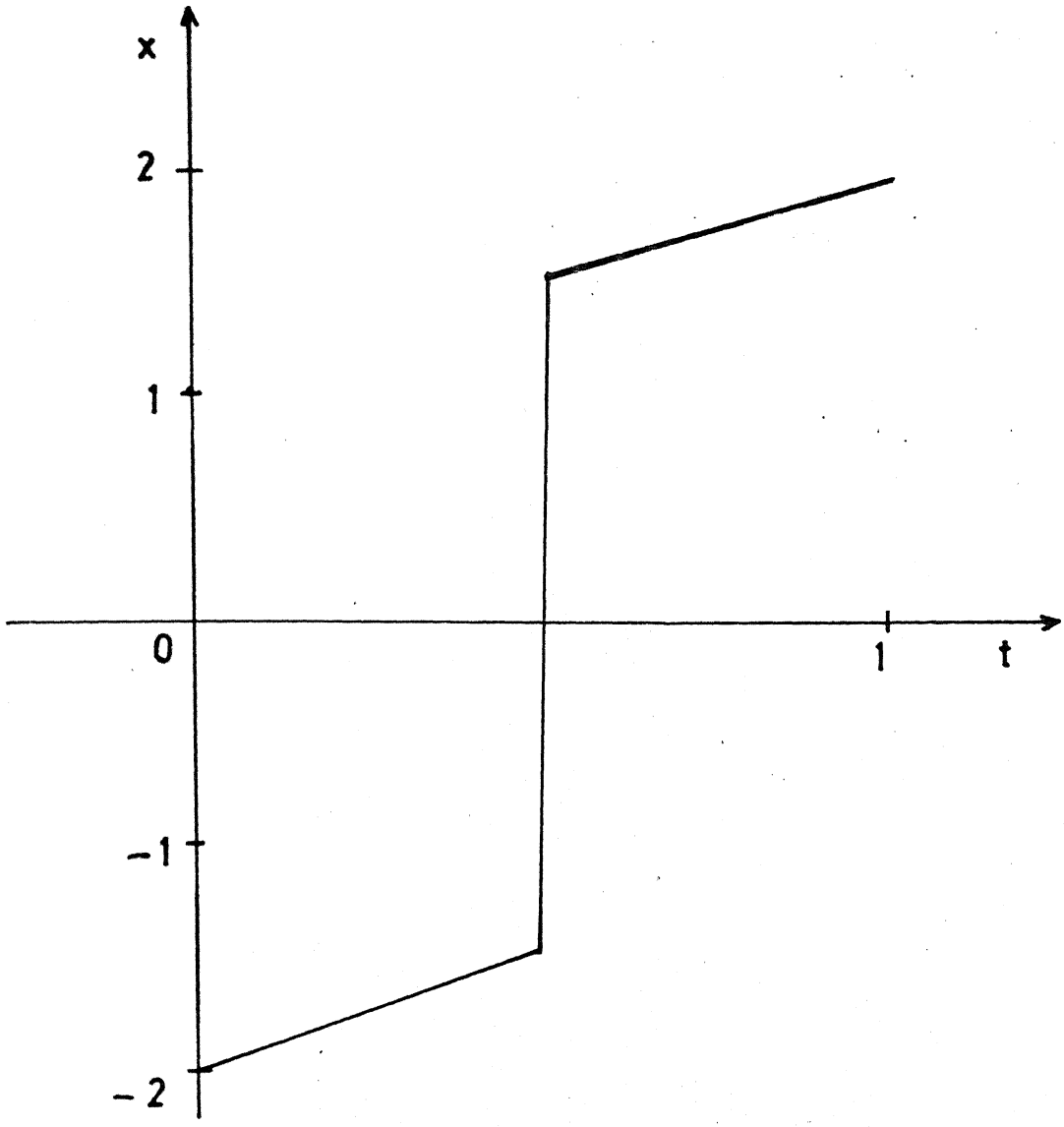


Figure 5.2

Numerical solution to case-3

$\varepsilon = 10^{-4}$, 4th iteration.

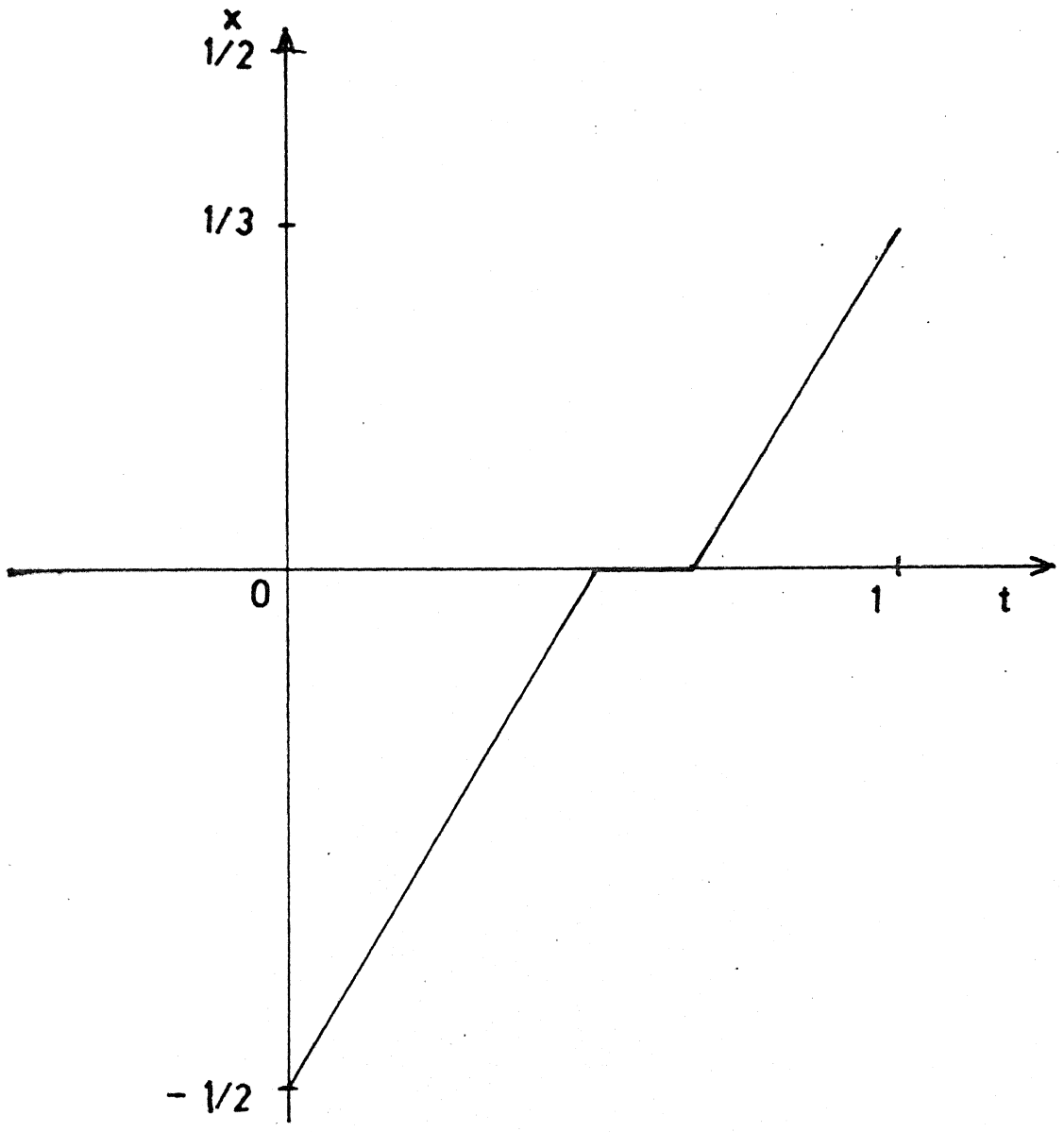


Figure 5.3

Numerical solution to case -4

$\varepsilon = 10^{-4}$, 4th iteration.

5.4 Discussions

As almost always it happens, the numerical solutions of non-linear problems via iterations is a lengthy process. We emphasize on the fact that the solution is to be computed after the non-linear system has been linearized and not before as e.g. in Keller [55]. For example, in getting a solution for Case-1, $\varepsilon = 10^{-2}$, we had to solve four linear systems obtained in the iteration steps. In spite of the long-winding process involved in the numerical solution, we prefer this approach because of the accuracy of the hybrid method suggested in Chapter 2, which we apply for computing the solutions of the linear systems.

When ε decreases to zero, the solution approaches the reduced solution except in the so called difficult regions (boundary layers, shock-layers and etc.). But the computation of numerical solution becomes more difficult when ε decreases to zero due to the increasing stiffness. However, it may be observed from the tables that the numerical solution becomes close to the asymptotic solution as ε decreases. This suggests the robustness of the proposed method. Again, neither by analytical nor

by asymptotical means it is possible to know the thickness of the layers. Since, numerically they are represented by the stretching sub-intervals, this method suggests an approximate thickness of the possible layers. But they are by no means conclusive. In Case-1, for example, we observe that a stretching sub-interval exists around the point $t = 1$, whereas there is no boundary layer at this point. This happened because the method sensed a possibility of such a layer. In general, if one needs finding out whether there is such a layer, one should also see whether there is a rapid change in the solution in the concerned region. From Figure 5.1 it can be seen that there exists no such change and hence, no boundary layer near $t = 1$.

We note that this method is also applicable to nonlinear stiff problems (Kreiss et al. [68], Sec.11). Stiff problems could also be tackled computationally via singular perturbations as e.g., in Aiken and Lapidus [3] and Flaherty and O'Malley [35].

We also remark that, once the linearization is over, the special methods of Chapter 4 can also be applied provided that they are applicable. For special non-linearities, such as quasi-linear systems

of singularly perturbed two-point boundary value problems, these methods would probably be applicable with advantages (specially the Cutting Point Technique).

Note that the asymptotic solution in Case-4 has five pieces, where as we observe from Figure 5.3, that the numerical solution does not show this. Instead we have only three pieces. This happens, because the asymptotic solution only smoothens the corners at $t = \frac{1}{2}$ and $\frac{2}{3}$; but, in fact they are not necessary, since x_L, x_M match at $t = \frac{1}{2}$ giving $x(1/2) = 0$ and also x_M and x_R match at $t = 2/3$ giving $x(2/3) = 0$. But since there are possibilities of corner layers, the grids were refined in the neighbourhoods of the points $t = 1/2$, $2/3$. This demonstrates the robustness of the proposed method.

Chapter 6

APPLICATIONS TO OPTIMAL CONTROL

6.1 Introduction

The origins of present day control theory can be traced back from the feedback theory in 1930's by Nyquist and Bode and the stability theory of nonlinear regulators by Lurie and Krasovski in 1940's. The two approaches differ in using frequency domain and time domain techniques. By the help of Pontryagin's Maximum Principle and Bellman's Dynamic Programming the last two and half a decade saw the curious merging of these two techniques. Kalman [52] synthesized these two methods analytically and Wonham [116] expanded the synthesis by the geometric methods. Within this period, curious applications of singular perturbations to control theory were also developed. In fact, long before that control engineers were using the simplified models for the practical problems they were facing. This new growing interest of applying singular perturbation to control problems [23] broke them the happiness that what they were doing was legitimate.

As it happened we did not go to singular perturbation in detail and we do not want to go to the singularly perturbed control problems in detail. In the last chapter we have suggested and studied some numerical methods to solve singularly perturbed systems of ordinary differential equations. Our aim in this chapter is to apply (some/all of) those methods to singularly perturbed optimal control problems. Details about these problems and other control problems can be had from the review papers of Kokotovic et al. [63], O'Malley [83],Kokotovic[60], Kokotovic et al. [62] and their references.

In this chapter we are interested in applying the numerical techniques developed in the preceding chapters to some optimal control problems. In § 6.2 the familiar linear state regulators are taken up; in § 6.3, the fixed end point optimal control problems and in § 6.4, we will consider non-linear regulators. § 6.5 gives the observations concerning the preceding sections.

6.2 Linear State Regulators

The problem here we consider is to select a continuous vector function $u(t,\varepsilon)$ of dimension r

that minimizes the scalar cost

$$J = \frac{1}{2}\left(z'(t_f,\varepsilon)\pi(\varepsilon)z(t_f,\varepsilon)\right) + \frac{1}{2}\int_{t_o}^{t_f}\left[z'(t,\varepsilon)Q(t,\varepsilon)z(t,\varepsilon) + u'(t,\varepsilon)R(t)u(t,\varepsilon)\right]dt \qquad (2.1a)$$

subject to the equations

$$\dot{x} = A_{11}(t,\varepsilon)x + A_{12}(t,\varepsilon)y + B_1(t,\varepsilon)u \qquad (2.1b)$$

$$\varepsilon\dot{y} = A_{21}(t,\varepsilon)x + A_{22}(t,\varepsilon)y + B_2(t,\varepsilon)u \qquad (2.1c)$$

with the end conditions

$$x(t_o,\varepsilon) = \alpha(\varepsilon), \qquad y(t_o,\varepsilon) = \beta(\varepsilon) \qquad (2.1d)$$

where $-\infty < t_o \le t \le t_f < \infty$, $\quad z = \begin{bmatrix} x \\ y \end{bmatrix}$,

x, y are continuously differentiable vector functions of dimensions m, n respectively; A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, R are smooth matrix functions of dimensions mxm, mxn, nxm, nxn, mxr, nxr, rxr respectively. Also,

$$Q = \begin{bmatrix} Q_{11}(t,\varepsilon) & Q_{12}(t,\varepsilon) \\ Q_{21}(t,\varepsilon) & Q_{22}(t,\varepsilon) \end{bmatrix}$$

$$\pi = \begin{bmatrix} \pi_{11}(\varepsilon) & \varepsilon\pi_{12}(\varepsilon) \\ \varepsilon\pi'_{12}(\varepsilon) & \varepsilon\pi_{22}(\varepsilon) \end{bmatrix}$$

where Q_{11}, Q_{12}, Q_{21}, Q_{22}, π_{11}, π_{12}, π_{22} are of dimensions mxm, mxn, nxm, nxn, mxm, mxn, nxn respectively, $\varepsilon > 0$ is a small parameter, '.' represents d/dt and (') denotes the transpose of a matrix; x, y are called the state variables and u, the control.

In physical problems, the small parameter $\varepsilon > 0$ often represents negligible time constants, small masses, moments of inertia, inductances, capacitances etc. For practical purposes, control engineers usually neglect the small parameter in the model they use for the specific problem and then carry on the analysis of the resulting reduced model. Of course, they have some reasons to justify the approximation; for, in the presence of ε, the dimensionality of the problem becomes large and also the differential equations guiding the states might be stiff. But sometimes, it happens that the reduced model is a 'bad' approximation to the original model. For some specific purposes, it is necessary to carry out the analysis without reducing the model (i.e.,

by taking $\varepsilon = 0$). To this end, we employ the numerical techniques developed in the preceding chapters.

It is well known that the problem (2.1) has a unique solution for fixed $\varepsilon > 0$, provided Q and π are symmetric positive semi definite matrices and R is a symmetric positive definite matrix. Basically, two different approaches are found in the literature that transform (2.1) to differential systems. One is via the Hamiltonian and the other, via the Riccati transforms. In general, for practical purposes, the systems of differential equations obtained by either of the above approaches, are solved by asymptotic techniques. We follow the Hamiltonian approach, due to its general applicability.

Introduce the Hamiltonian

$$H(x,y,p,q,u,t,\varepsilon) = \frac{1}{2}(z'Qz + u'Ru) + p'(A_{11}x + A_{12}y + B_1u) + q'(A_{21}x + A_{22}y + B_2u) \tag{2.2}$$

for differentiable m,n dimensional vector functions $p(t,\varepsilon)$, $q(t,\varepsilon)$ (called costate variables) that satisfy the equations

$$\dot{p} = -\partial H/\partial x \quad , \qquad \varepsilon \dot{q} = -\partial H/\partial y \qquad (2.3a)$$

with the terminal conditions

$$p(t_f,\varepsilon) = \partial\lambda/\partial x]t = t_f, \quad \varepsilon q(t_f,\varepsilon) = \partial\lambda/\partial y] = t_f \qquad (2.3b)$$

where $\lambda(\varepsilon) = \frac{1}{2} z'(t_f,\varepsilon)\pi(\varepsilon)z(t_f,\varepsilon)$, the expression which does not come under the integral sign in (2.1a). It is known [Pontryagin's theorem, i.e., Thm 6.4 of Chapter 1] that the minimization of the cost functional J in (2.1a) is equivalent to the minimization of the Hamiltonian H in (2.2). Thus $\partial H/\partial u = 0$ gives

$$u = -R^{-1}(B_1'p + B_2'q). \qquad (2.4)$$

Again, since $\partial^2 H/\partial u^2 = R$ which is assumed to be positive definite, H is minimized by (2.4) and in turn J. Hence u is determined by (2.4) provided that $x,y,\ p,q$ are determined as well. x, y, p, q are solutions of the following system (Q being symmetric, $Q_{21} = Q_{12}'$).

$$\dot{x} = A_{11}x + A_{12}y - B_1R^{-1}B_1'p - B_1R^{-1}B_2'q \qquad (2.5a)$$

$$\varepsilon\dot{y} = A_{21}x + A_{22}y - B_2R^{-1}B_1'p - B_2R^{-1}B_2'q \qquad (2.5b)$$

$$\dot{p} = -Q_{11}x - Q_{12}y - A_{11}'p - A_{21}'q \qquad (2.5c)$$

with the end conditions :

$$x(t_o,\varepsilon) = \alpha(\varepsilon), \qquad y(t_o,\varepsilon) = \beta(\varepsilon) \tag{2.5e}$$

$$p(t_f,\varepsilon) = \pi_{11}x(t_f,\varepsilon) + \varepsilon\,\pi_{12}y(t_f,\varepsilon) \tag{2.5f}$$

$$q(t_f,\varepsilon) = \pi'_{12}x(t_f,\varepsilon) + \pi_{22}y(t_f,\varepsilon) \tag{2.5g}$$

The two-point boundary-value system (2.5) is a linear homogeneous singularly-perturbed TPBVS; which can be solved by any of the methods suggested in the preceding chapters; and the control u can be determined by using equation (2.4).

Below, we take two familiar examples to illustrate the methods. Solutions to both of them have been studied earlier [61,81].

The hybrid method is expected to work well due to the underlying analysis. So here we will demonstrate the other two methods, namely, Boundary Value Technique and Cutting Point Technique.

Example 1 ([81])

It is required to select the control $u(t,\varepsilon)$ to minimize the cost functional

$$J = \tfrac{1}{2}[x^2(1) + \varepsilon y^2(1)] + \tfrac{1}{2}\int_o^1 [x^2(t) + u^2(t)]dt \tag{2.6a}$$

$$\dot{x} = x \quad , \quad x(0) = 1.3708 \qquad (2.6b)$$

$$\varepsilon\dot{y} = -x - y + u \quad , \quad y(0) = 5.9420 \qquad (2.6c)$$

Comparing with the equations (2.1) we see that

$$\pi = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix} , \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} ,$$

$R = 1$, $A_{11} = 1$, $A_{12} = 0$, $A_{21} = A_{22} = -1$, $B_1 = 0$ and $B_2 = 1$. Hence the Hamiltonian reads as follows.

$$H(x,y,p,q,u,t,\varepsilon) = \frac{1}{2}(x^2+u^2) + px + q(-x-y+u).$$

Minimization of H gives us the following TPBVP :

$$\dot{x} = x \quad , \quad x(0) = 1.3608 \qquad (2.7a)$$

$$\varepsilon\dot{y} = -x - y - q \quad , \quad y(0) = 5.9420 \qquad (2.7b)$$

$$\dot{p} = -x - p + q \quad , \quad p(1) = x(1) \qquad (2.7c)$$

$$\varepsilon\dot{q} = q \quad , \quad q(1) = y(1) \qquad (2.7d)$$

$$0 = u + q$$

We want to apply Boundary Value Technique. (2.7a) gives the solution of $x(t)$ directly, i.e.,

$$x(t) = 1.3708 \exp(t) \qquad (2.8a)$$

Eliminating q from (2.7b) - (2.7d), we obtain :

$$\varepsilon \ddot{y} = y + (1-\varepsilon)x \qquad (2.8b)$$

$$\dot{p} + p + 2x - y - \varepsilon \dot{y} = 0 \qquad (2.8c)$$

$$u = x + y + \varepsilon \dot{y} \qquad (2.8d)$$

The end conditions are :

$$p(1) = 1.3708\ e \qquad (2.8e)$$

$$\varepsilon \dot{y}(1) = -1.3708\ e \qquad (2.8f)$$

$$y(0) = 5.9420. \qquad (2.8g)$$

Letting $p = \sum_{i=0}^{\infty} p_i \varepsilon^i$, $y = \sum_{i=0}^{\infty} y_i \varepsilon^i$ in (2.8) and comparing like powers of ε, we determine p, y upto $O(\varepsilon)$ terms, with $t_1 = 0.07$, $t_2 = 0.96$; we have (denoting these values with a bar),

$$\bar{p}(t_1) = 2.2065 \quad , \quad \bar{p}(t_2) = 9.5137 \qquad (2.9a)$$

$$\bar{y}(t_1) = 0.2134 \times 10^{-2} \quad , \quad \bar{y}(t_2) = 1.5327 \qquad (2.9b)$$

Then we solve the following boundary value problems :

($\tau = t/\varepsilon$, $\sigma = (1-t)/\varepsilon$)

$$dp/d\tau + \varepsilon(p - y - dy/d\tau + 2x) = 0 \qquad (2.10a)$$

$$d^2y/d\tau^2 - \varepsilon(y + (1-\varepsilon)x) = 0 \qquad (2.10b)$$

$$y(0) = 5.9420 \tag{2.10c}$$

$$y(t_1/\varepsilon) = \bar{y}(t_1), \quad p(t_1/\varepsilon) = \bar{p}(t_1) \tag{2.10d}$$

for $0 \leq \tau \leq t_1/\varepsilon$.

Next,

$$-dp/d\sigma + \varepsilon(p - y + dy/d\sigma + 2x) = 0 \tag{2.11a}$$

$$d^2y/d\sigma^2 + (y + (1-\varepsilon)x) = 0 \tag{2.11b}$$

$$p(0) = 1.3708e, \quad y((1-t_2)/\varepsilon) = \bar{y}(t_2) \tag{2.11c}$$

$$(1+\varepsilon)y(0) + \varepsilon dy(0)/d\sigma + 1.3708e = 0 \tag{2.11d}$$

for $0 \leq \sigma \leq (1-t_2)/\varepsilon$.

Finally,

$$p + \dot{p} - y - \varepsilon\dot{y} + 2x = 0 \tag{2.12a}$$

$$\varepsilon\ddot{y} - y - (1-\varepsilon)x = 0 \tag{2.12b}$$

$$p(t_1) = \bar{p}(t_1), \quad y(t_1) = \bar{y}(t_1) \tag{2.12c}$$

$$y(t_2) = \bar{\bar{y}}(t_2) \tag{2.12d}$$

for $t_1 \leq t \leq t_2$.

We solve (2.10 - 2.12) with the tolerence limit equal to 0.03.

The exact solution of $u(t)$ is approximately (neglecting the exponentially small quantities) given by

$$u(t) = \frac{2}{3}\frac{1.3708e}{1+\varepsilon} \cdot \exp((t-1)/\varepsilon) \qquad (2.13)$$

The comparisons of the numerical solution (by Boundary Value Technique) and the exact solution approximated as in (2.13) are summarized in Table 6.1.

Note that the method worked even if we do not choose t_1, t_2 afresh for different ε's.

Example 2 ([61])

Consider a speed control problem for a small D/C motor. The state equations are :

$$dw/dt = (D/G)i$$

$$\varepsilon L\, di/dt = -Cw - R_a\, i + v$$

where w, i and v are speed, current and voltage deviations from their nominal values 400 rad/s, 0.25A and 4.3V. The constants are $R_a = 7.9\,\Omega$, $D/G = 1.8654$, $L = 0.0136H$, $C = 0.0246$ Vs/rad. The cost to be minimized is :

$$J = \frac{0.01\, w^2(1)}{2} + \frac{1}{2}\int_0^1 (w^2+400i^2+30v^2)dt.$$

Here $\pi = \begin{bmatrix} 0.01 & 0 \\ 0 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 0 \\ 0 & 400 \end{bmatrix}$

$R = 30$, v is the control, w and i are the state variables. $A_{11} = 0$, $A_{12} = D/G$, $A_{21} = -C/L$, $A_{22} = -R_a/L$, $B_1 = 0$, $B_2 = 1/L$.

Thus, the Hamiltonian is :

$$H(w, i, p, q, v, t, \varepsilon) = \frac{1}{2}(w^2 + 400i^2 + 30v^2) + (D/G)pi + q(-\frac{C}{L}w - \frac{R_a}{L}i + \frac{1}{L}v).$$

Minimization of H gives us the following TPBVP : (Retaining v instead of q, as $H_v = 0$ implies $q = -30L\,v$);

$$dw/dt = (D/G)i \qquad (2.14a)$$

$$\varepsilon\, di/dt = -(C/L)w - (R_a/L)i + (1/L)v \qquad (2.14b)$$

$$dp/dt = -w - 30\,Cv \qquad (2.14c)$$

$$\varepsilon\, dv/dt = (40/3L)i + (D/30LG)p + (R_a/L)v \qquad (2.14d)$$

with the end conditions :

$$w(0) = 400, \quad i(0) = 0.25 \qquad (2.14e)$$

$$p(1) = 0.01w(1), \quad v(1) = 0. \qquad (2.14f)$$

We apply Boundary Value Technique to solve (2.14) (with tolerence limit = 0.03). The reduced solution for the control v is given by

$$v(t) = -38.734094 \exp(at) + 42.726941 \exp(at) \qquad (2.15)$$

with $a = 0.49055$.

Comparisons of the numerical solution with the reduced solution (2.15) is given in Table 6.1. In the table the same terminologies are followed; Mad, Mid .

The solution of the linear regulators in general, contains boundary layers, whence the left boundary layer is stable (in 't') and the right boundary layer is stable in reverse time (i.e., in '-t'). For the occurrence of boundary layers, we assume that the system (2.1) is boundary layer observable and boundary layer controllable [61]. Further assumptions that we need are : x, y, p, q are twice continuously differentiable vector functions and the conditions needed for the existence and uniqueness of the solution to this problem. Under these assumptions we have obtained a solution to the problem using τ- and σ-scaled equations.

The computed results for Example 1 show that the solution obtained by the proposed method (Boundary Value Technique) approximate the original solution well as the small positive parameter ε is made smaller, which is a necessary feature for numerical methods designed for singular perturbation problems. From the second example, we observe that the solution by the proposed method (Boundary Value Technique) does approach the reduced solution as ε approaches zero. (In the second example, we have taken the reduced solution for comparison purposes, for the stiffness involved in the problem).

It has been observed that the method works well on linear optimal control problems. The applicability of the present methods to fixed-end point problems and non-linear regulators come as possible generalizations. In later sections, these possibilities will be investigated.

Table 6.1

Summary of the Comparison Results for the Examples

Ex - 1			Ex - 2		
ε			ε		
1E-1	Mad	0.711E-3	0.5	Mad	0.099561
	Mid	0.35E-6		Mid	0.028142
1E-2	Mad	0.225E-3	0.1	Mad	0.10271
	Mid	0.1E-6		Mid	0.000928
1E-3	Mad	0.155E-3	0.05	Mad	0.013471
	Mid	0.0E-1		Mid	0.000021
1E-4	Mad	0.478E-9	0.01	Mad	0.000173
	Mid	0.0E-1		Mid	0.000003
1E-5	Mad	0.248E-11	0.001	Mad	0.000002
	Mid	0.0E-1		Mid	0.000000

6.3 Singularly Perturbed Linear Fixed End-Point Optimal Control and Cheap Control

In § 6.2 we have considered the linear regulators and demonstrated the efficiency of the numerical method (Boundary value technique) for such problems. A natural extension is the fixed end-point problems, i.e., when the end-points $x(t_o)$, $x(t_f)$, $y(t_o)$, $y(t_f)$ of the (state) trajectories are fixed. (see equation - (2.1d)). Precisely, the problem is to select control $u(t,\varepsilon)$ that minimizes the cost (integral cost)

$$J = \frac{1}{2} \int_{t_o}^{t_f} (z'z + u'Ru)dt \tag{3.1a}$$

subject to the state

$$\dot{x} = A_{11}x + A_{12}y + B_1y \tag{3.1b}$$

$$\varepsilon\dot{y} = A_{21}x + A_{22}y + B_2y \tag{3.1c}$$

with the boundary conditions

$$x(t_o,\varepsilon) = x^o, \quad y(t_o,\varepsilon) = y^o \tag{3.1d}$$

$$x(t_f,\varepsilon) = x^f, \quad y(t_f,\varepsilon) = y^f \tag{3.1e}$$

where $z = C_1(t)x + C_2(t)y$, C_1, C_2 being given matrices of functions of t; and other things

reading the same as defined in § 6.2.

Applying Pontryagin's Principle (Theorem 6.4 of Chapter 1) to (3.1) we find that the optimal trajectory, (x,y) the co-states (p,q) and the control u are given as

$$u = -R^{-1}(B_1'p + B_2'q) \tag{3.2a}$$

$$\dot{x} = A_{11}x + A_{12}y - S_{11}p - S_{12}q \tag{3.2b}$$

$$\varepsilon\dot{y} = A_{21}x + A_{22}y - S_{12}'p - S_{22}q \tag{3.2c}$$

$$\dot{p} = -C_1'C_1x - C_1'C_2y - A_{11}'p - A_{12}'q \tag{3.2d}$$

$$\varepsilon\dot{q} = -C_2'C_1x - C_2'C_2y - A_{21}'p - A_{22}'q \tag{3.2e}$$

with boundary conditions (3.1d - 3.1e) and

$$S_{ij} = B_iR^{-1}B_j' \quad , \qquad i,j = 1,2.$$

For the applications of this problem one is referred to Sannuti [102], Khalil [59], Hadlock [40]. It is known that this problem is not robust to singular perturbations. In fact, the left-boundary layer is stable in time whereas the right boundary layer is stable in reverse-time (negative time) [37, 40, 102]. This is usually solved either by a Riccati equations approach or by the method of Matched Asymptotic expansions. The difficulties associated with

these approaches are with finding positive and negative definite solutions of the Riccati equations [114] or with 'matching' in a hypothetical domain. The numerical methods developed in the previous chapters apply well to this case. The TPBVP contained in (3.2) with the boundary conditions (3.1d - 3.1e) can be solved in a similar manner as the corresponding problem in § 6.2.

Now we consider another related linear control problem, namely Cheap Control. Consider a cost functional (to be minimized)

$$J(\varepsilon) = \frac{1}{2} \int_{t_o}^{t_f} (x'Qx + \varepsilon^2 u'Ru)dt \qquad (3.3a)$$

Q, being symmetric positive semidefinite and R, symmetric positive definite smooth matrix functions respectively. The control constraints being in the time-varying system form

$$\dot{x} = Ax + Bu \qquad t_o < t < t_f \qquad (3.3b)$$

with an initial condition

$$x(t_o) = x^o \qquad (3.3c)$$

where $A(t)$, $B(t)$ are smooth matrix functions of dimensions nxn and nxr ; x and u are column

vectors of dimensions n and r respectively; $\varepsilon > 0$ is a small parameter. (Note that the appearance of ε is in the cost and not in the states). Classical control theory implies that [52] the optimal control

$$u = -\varepsilon^{-2} R^{-1} B' p \tag{3.4a}$$

where the co-state variable p satisfies

$$\dot{p} = -Qx - A'p \tag{3.4b}$$

and the end-condition

$$p(t_f) = 0. \tag{3.4c}$$

Eliminating u, from (3.3 - 3.4), we have the following singularly perturbed TPBVP :

$$\varepsilon^2 \dot{x} = \varepsilon^2 Ax - BR^{-1} B' p \tag{3.5a}$$

$$\dot{p} = -Qx - A'p \tag{3.5b}$$

$$x(t_o) = x^o, \qquad p(t_f) = 0. \tag{3.5c}$$

which can be solved numerically. However, as in O'Malley and Jameson [87] we will consider a special subclass of such cheap control problems.

<u>Definition 3.1</u> Define 'Case-L' by requiring that

$B_j^T Q B_j = 0 \qquad j = 0,1,\ldots,L-2$

$B_{L-1}^T Q B_{L-1} > 0$

on $[t_o, t_f]$ with

$B_j = AB_{j-1} - \dot{B}_{j-1}$, $j \geq 1$; $\qquad B_o = B$.

The special subclass refers to the problems that belong to Case-L, for some L.
It is shown in [87] that for Case-L, the boundary layer variables (τ and σ previously) must be chosen by

$$\tau = (t-t_o)/(\varepsilon)^{1/L} \quad , \quad \sigma = (t_f-t)/(\varepsilon)^{1/L} \qquad (3.6)$$

Hence, when applying the special methods (i.e., Boundary value technique and cutting point technique), the boundary layers must be stretched according to (3.6).

Note that we need not transform the problems into an $n+Lr$ - dimensional problem as in [87] for having a numerical solution. This suggests an obvious advantage of the suggested numerical approaches over those employed in the afore-cited reference. For more details, one is advised to see appropriate references in Kokotovic [60].

6.4 Nonlinear State Regulators

In this section we consider an optimization problem of Lagrange type where the system equations are stiff and the system can be expressed in terms of a singular perturbation parameter. Specifically, we consider the problem of minimizing the scalar cost

$$J = \pi(x(t_f), y(t_f), \varepsilon) + \int_{t_o}^{t_f} \varphi(x,y,u,t,\varepsilon)dt \quad (4.1a)$$

subject to the state constraints

$$\dot{x} = f(x,y,u,t,\varepsilon) \quad (4.1b)$$

$$\varepsilon\dot{y} = g(x,y,u,t,\varepsilon) \quad (4.1c)$$

$$x(t_o) = x^o, \qquad y(t_o) = y^o \quad (4.1d)$$

where $x(t,\varepsilon)$ and $y(t,\varepsilon)$ are smooth state variables of dimensions m and n respectively, $u(t,\varepsilon)$ is the control vector of dimension r, the real number ε is a small positive parameter and '$\cdot$' denotes d/dt. It is desired to select the control u to minimize J and thereby determine the state trajectory.

In physical problems, the small positive parameter ε represents inductances, capacitances,

moments of inertia, and small masses, which control engineers often neglect. Instead of taking the problem (4.1), they consider the reduced problem (i.e. by letting $\varepsilon = 0$) because of the difficulties that arise from the high dimensionality and the stiffness of the system (4.1). But in many cases, the reduced model becomes a 'bad' approximation to the original model and it loses many of the important properties.

When ε is close to zero, the problems are such that near $t = t_o$ and $t = t_f$ there exists no uniformly valid solution. In other words, the solution of the original problem does not converge uniformly to the solution of the reduced one as $\varepsilon \longrightarrow 0$. Such a situation is commonly referred to as the occurrence of a boundary layer at $t = t_o$ and at $t = t_f$. Mostly, the solutions to these problems have been constructed with the help of matched asymptotic expansions (cf. Kokotovic [60]). Then the solutions of the boundary-layer equations and that of the reduced equations are combined in some way to give an approximate solution over the entire interval. As usual, define the Hamiltonian H with the help of co-state variables p and q

of dimensions m and n respectively :

$$H(x, y, p, q, u, t, \varepsilon) = \varphi + p'f + q'g ,$$

where (') represents transpose of a vector or matrix. The necessary conditions for optimality follow from Pontryagin's principles [Thm. 6.4 of Chapter 1]. In addition to (4.1b) - (4.1d), we have

$$\dot{p} = -\partial H/\partial x , \quad \varepsilon \dot{q} = -\partial H/\partial y \tag{4.2a}$$

$$p(t_f,\varepsilon) = \partial\pi(t_f,\varepsilon)/\partial x, \quad \varepsilon q(t_f,\varepsilon) = \partial\pi(t_f,\varepsilon)/\partial y \tag{4.2b}$$

$$\partial H/\partial u = 0 , \tag{4.2c}$$

$\partial^2 H/\partial u^2 > 0$ is sufficient for the minimization of J . Now, our task is to solve (4.1b) - (4.2b) numerically and determine u from (4.2c). Note that if we want to use the Hybrid method we have to follow the linearization suggested in Chapter 5. On the other hand, iterations can be avoided by using the special methods developed in Chapter 4. In the following, Cutting Point Technique is modified to solve the above problem class. Boundary Value Technique can also be modified similarly. However, if the problem has a turning point, one is advised to follow the Hybrid method strictly.

For the purpose of solving the TPBVP (4.1b-4.2b),

we assume the following :

(A-1) The strong Legendre-Clebsch condition $\partial^2 H/\partial u^2 > 0$ holds.

(A-2) $\pi(x,y,\varepsilon) = \theta(x, \varepsilon y, \varepsilon)$ for some functions $\theta(.,.,.)$.

(A-3) The reduced problem of (4.1b) - (4.2b) has a unique solution (with appropriate boundary conditions)

(A-4) The boundary layers are locally stable (with dichotomies, i.e., including reverse-stability near $t = t_f$) and there is no turning point except in the neighbourhoods of t_o and t_f (stretching factor = ε).

Assumption A-1 guarantees that $H_u = 0$ in (4.2c) can be solved uniquely to yield $u = \eta(x,y,p,q,t,\varepsilon)$ (see Thm. 6.5 of Chapter 1); A-2 tells that π is a slowly varying function of the fast state variable y; A-3 will help in finding out values of x,y,p,q at the points t_1 and t_2 (necessary for the Cutting Point Technique). Assumption A-4 forms the basis of scaling the equations by τ and σ. With these assumptions (of which A-3 and A-4 are exclusively necessary for the special method : i.e., Cutting Point Technique), we proceed to describe the technique in a stepwise manner, as follows :

Step 1

Choose $t_1, t_2 \in [t_o, t_f]$ close to t_o and t_f respectively. We expect that the boundary layers lie within the intervals $[t_o, t_1]$ and $[t_2, t_f]$, (as experience suggests we let the thickness of the boundary layers be $O(|\varepsilon \ln \varepsilon|)$). With this expectation, find the values of x, y, p, q at t_1 and t_2 from the reduced equations corresponding to (4.1b - 4.2b), neglecting the boundary conditions for $y(t_o)$ and $q(t_f)$; i.e., we evaluate $x(t_1)$, $y(t_1)$ etc. from the following either by analytical or by numerical means :

$$\dot{x} = f, \quad \dot{p} = -\partial H/\partial x, \quad g = 0 = \partial H/\partial y \tag{4.3a}$$

$$x(t_o) = x^o, \quad p(t_f) = \partial \pi(t_f, 0)/\partial x \tag{4.3b}$$

Denote $x(t_1) = x^1$, $x(t_2) = x^2$, etc.

Step 2

In this step, our aim is to get an approximate solution to the problem (4.1b) - (4.2b). To this end, we let

$$\tau = (t - t_o)/\varepsilon \quad , \quad \sigma = (t_f - t)/\varepsilon \tag{4.4a}$$

$$\tau_1 = (t_1 - t_o)/\varepsilon \quad , \quad \sigma_2 = (t_f - t_2)/\varepsilon \tag{4.4b}$$

and form the τ-scaled and the σ-scaled equation of

(4.1b) - (4.2b), using the transformations (4.4a). We supply the τ-scaled, original and σ-scaled equations with the boundary conditions as in the following. . . .

τ-scaled	Original	σ-scaled
x,y at t_o	x,y at t_1	x,y at t_2
p,q at t_1	p,q at t_2	p,q at t_f .

Let $X = \langle x,y,p,q \rangle$ denote the stacked $(2m+2n)$-vector $(x',y',p',q')'$ and (note the use of $\langle .,.,.,. \rangle$)

$$\text{let} \quad A = \begin{bmatrix} I_{m+n} & O_{m+n} \\ O_{m+n} & O_{m+n} \end{bmatrix}, \quad B = \begin{bmatrix} O_{m+n} & O_{m+n} \\ O_{m+n} & I_{m+n} \end{bmatrix}$$

where I_k is the identity matrix of order k and O_k is the square zero matrix of order k, and proceed as follows :

Solve (4.5) for $\tau \in [0, \tau_1]$:

$$\dot{X} = \langle f, g, -\partial H/\partial x, -\partial H/\partial y \rangle \qquad (4.5a)$$

$$AX(0) + BX(\tau_1) = \langle x^o, y^o, p^1, q^1 \rangle \qquad (4.5b)$$

Solve (4.6) for $t \in [t_1, t_2]$:

$$\dot{X} = \langle f, \varepsilon^{-1}g, -\partial H/\partial x, -\varepsilon^{-1}\partial H/\partial y \rangle \qquad (4.6a)$$

$$AX(t_1) + BX(t_2) = \langle x^1, y^1, p^2, q^2 \rangle \qquad (4.6b)$$

Finally, solve (4.7) for $\sigma \in [0,\sigma_2]$:

$$\dot{X} = \langle -\varepsilon f, -g, \varepsilon \partial H/\partial x, \partial H/\partial y \rangle \qquad (4.7a)$$

$$BX(0) + AX(\sigma_2) = \langle x^2, y^2, \partial\pi(t_f,\varepsilon)/\partial x, \varepsilon^{-1}\partial\pi(t_f,\varepsilon)/\partial y \rangle \qquad (4.7b)$$

Next, we combine the solutions of (4.5 - 4.7) using the inverse transformations of (4.4a) :

$t = \varepsilon\tau + t_o$ on $[0, \tau_1]$ and

$t = t_f - \varepsilon\sigma$ on $[0,\sigma_2]$.

Step 3

In this step, we describe briefly how to solve the nonlinear TPBVP's obtained above. Here we apply centered difference scheme. Our problems are of the following form :

$$NX \equiv \dot{X} - F(t,X) = 0, \quad AX(a) + BX(b) = C \qquad (4.8)$$

where N is a nonlinear operator and C is a $2(m+n)$-dimensional constant vector. To ensure the convergence of the Newton's iteration scheme (described below) we assume that F satisfies the following conditions :

(i) $\dfrac{\partial F_i}{\partial X_j}(t,X)$ is continuous for, $i,j = 1,2,..,2m+2n$.

$$\Psi(Z) \equiv \begin{bmatrix} A\chi_o + B\chi_k - C \\ hN_h\chi_1 \\ hN_h\chi_k \end{bmatrix} = 0. \qquad (4.11)$$

Now, with some initial guess Z_o (we might take the reduced solution) we compute the sequence Z_s, $s = 0,1,\ldots$ by the following :

$$Z_{s+1} = Z_s + \bar{Z}_s \qquad (4.12a)$$

where $\bar{Z}_s$ is the solution of the linear algebraic system

$$\frac{\partial \Psi}{\partial Z}(Z_s)\,\bar{Z}_s + \Psi(Z_s) = 0 \qquad (4.12b)$$

Note that the error in this procedure is $O(h^2)$. (cf. [55, § 3.3]).

Step 4

In this step, our aim is to confirm whether our choice of t_1 and t_2 is acceptable, or whether we need to seek a better choice. To this end, we fix up a tolerence limit on the difference between the derivatives (given below) computed from the three different regions $[t_o,t_1]$, $[t_1,t_2]$, $[t_2,t_f]$ at the points t_1, t_2. So, we compute $\dot{\chi}(t_1)$ from (4.5 - 4.6) and denote them by $\dot{\chi}_-(t_1)$ and $\dot{\chi}_+(t_1)$

respectively; we also compute $\dot{X}(t_2)$ from (4.6 - 4.7) to denote them by $\dot{X}_-(t_2)$ and $\dot{X}_+(t_2)$ respectively. We then suggest the following criterion which is to be satisfied by our computed solution :

$$||\dot{X}_+(t_j) - \dot{X}_-(t_j)|| < \delta , \quad j = 1,2 \qquad (4.13)$$

where $\delta > 0$ is some prescribed tolerence limit. If (4.13) is satisfied, we stop, else, we choose different t_1 greater than, and t_2 less than that previously chosen and repeat the steps 1-4.

In the following, we take two examples to illustrate the method. The first example is a quasilinear one taken from Sannuti [103] and the second is from Freedman and Granoff [36] (also Ardema [4]) which is a nonlinear.

Example 3

Consider the problem of minimizing

$$J = \frac{1}{2} \int_0^1 (x^2 + u^2)dt$$

with the state equations

$$\dot{x} = y + \varepsilon x^2 , \qquad x(0) = x^o = 11.45$$

$$\varepsilon\dot{y} = - x - y + u, \quad y(0) = y^o = 1.985 .$$

The Hamiltonian is

$$H(x,y,p,q,u,t,\varepsilon) = -\frac{1}{2}(x^2+u^2) + p(y+\varepsilon x^2) + q(-x-y+u).$$

The TPBVP corresponding to (4.1b) - (4.2b) is :

$$\dot{x} = y + \varepsilon x^2 , \qquad x(0) = x^0 \qquad (4.14a)$$

$$\varepsilon\dot{y} = -x - y + q , \qquad y(0) = y^0 \qquad (4.14b)$$

$$\dot{p} = -x - 2\varepsilon xp + q, \qquad p(1) = 0 \qquad (4.14c)$$

$$\dot{q} = -p + q , \qquad q(1) = 0 \qquad (4.14d)$$

$$u = q . \qquad (4.15)$$

The reduced solution of (4.14) without the conditions $y(0) = y^0$ and $q(1) = 0$ is :

$$x = x^0(1 - t/2) , \quad y = -x^0/2 , \qquad (4.16a)$$

$$p = x^0(1 - t)/2 , \quad q = x^0(1 - t)/2. \qquad (4.16b)$$

We choose t_1 and t_2 corresponding to various ε's as follows :

ε	0.1	0.001	0.00001
t_1	0.05	0.005	0.00005
t_2	0.05	0.005	0.00005

Next we compute $x(t_j)$ etc. from (4.16) with

these choices of t_j, $j = 1,2$; and denote these by $s(t_j) = s^j$, $s \in \{x,y,p,q\}$, $j = 1,2$. The problems in step 2 to be solved are the following :

Denote $\tau = t/\varepsilon$, $\sigma = (1-t)/\varepsilon$, $\tau_1 = t_1/\varepsilon$ and $\sigma_2 = (1-t_2)/\varepsilon$,

The equations corresponding to (4.5) are :

$$dx/d\tau = \varepsilon^2 x^2 + \varepsilon y \quad , \quad x(0) = x^o \qquad (4.17a)$$

$$dy/d\tau = - x - y + q \quad , \quad y(0) = y^o \qquad (4.17b)$$

$$dp/d\tau = - \varepsilon x - 2\varepsilon^2 xp + q \quad , \quad p(\tau_1) = p^1 \qquad (4.17c)$$

$$dq/d\tau = - p + q \quad , \quad q(\tau_1) = q^1 \qquad (4.17d)$$

Similarly the equations corresponding to (4.6 - 4.7) might be formed. Note that all τ-scaled, original and σ-scaled equations corresponding to (4.5) - (4.7) satisfy four assumptions in step 3. Finally, we combine the results of (4.17). &c. , to get a solution for (4.14), using the transformations $t = \varepsilon\tau$ or $t = 1 - \varepsilon\sigma$ in the appropriate domains. Since in this modification we have followed the conventional approach of taking finite difference schemes first and apply linearization afterwards, we give the full numerical results in the following tables instead of

just the comparison results as we were following previously. The numerical solutions are given in Tables (6.2) - (6.4).

Example 4

Select the controls u and v to minimize the terminal cost

$J = x(1,\varepsilon) + \frac{1}{2}\varepsilon y(1,\varepsilon)$ subject to the state constraints

$$\dot{x} = y + u^2 - tv + \varepsilon(x+u) = f$$

$$\varepsilon\dot{y} = -y - tu + 2v^2 + \varepsilon^2(x+y^2) = g$$

$x(0,\varepsilon) = x^0 = 28.239, \quad y(0,\varepsilon) = y^0 = 5.984.$

Considering the corresponding Hamiltonian (with $\pi = J$)

$$H = pf + qg ,$$

we have the following TPBVP :

$\dot{x} = f, \quad \varepsilon\dot{y} = g, \qquad x(0) = x^0, \quad y(0) = y^0$

$\dot{p} = -\varepsilon p - \varepsilon^2 q , \qquad p(1,\varepsilon) = 1$

$\varepsilon\dot{q} = -p + q - 2\varepsilon^2 yq, \qquad q(1,\varepsilon) = \frac{1}{2} ,$

$u = tq/2(p-\varepsilon) , \qquad v = tp/4q,$

whose reduced solution is :

$$x = x_o - \frac{3}{8} t^3, \quad y = -\frac{3}{8} t^2, \quad p = q = 1.$$

We choose t_1 and t_2 for corresponding ε's as follows :

ε	0.1	0.001	0.00001
t_1	0.05	0.005	0.0005
t_2	0.05	0.005	0.0005

and compute $\omega(t_j) = \omega^j$ and solve the following systems : (see equations (4.5) - (4.7)) for $\tau \in [0, \tau_1]$:

$$dx/d\tau = \varepsilon(y + u^2 - tv) + \varepsilon^2(x + u)$$

$$dy/d\tau = -(y + tu) + 2v^2 + \varepsilon^2(x + y^2)$$

$$dp/d\tau = -\varepsilon^2(p + \varepsilon q)$$

$$dq/d\tau = -p + q - 2\varepsilon^2 yq$$

$$x(0) = x^o, \quad y(0) = y^o, \quad p(\tau_1) = p^1, \quad q(\tau_1) = q^1;$$

for $t \in [t_1, t_2]$:

$$\dot{x} = y + u^2 - tv + \varepsilon(x + u)$$

$$\varepsilon\dot{y} = -(y + tu - 2v^2) + \varepsilon^2(x + y^2)$$

$$\dot{p} = -\varepsilon(p + \varepsilon q)$$

$$\varepsilon\dot{q} = -p + q - 2\varepsilon^2 yq.$$

$$x(t_1) = x^1, \quad y(t_1) = y^1, \quad p(t_2) = p^2, \quad q(t_2) = q^2;$$

finally, for $\sigma \in [0, \sigma_2]$:

$$dx/d\sigma = -\varepsilon(y + u^2 - tv) - \varepsilon^2(x + u)$$
$$dy/d\sigma = y + tu - 2v^2 - \varepsilon^2(x + y^2)$$
$$dp/d\sigma = \varepsilon^2(p + \varepsilon q)$$
$$dq/d\sigma = p - q + 2\varepsilon^2 yq$$
$$x(\sigma_2) = x^2, \quad y(\sigma_2) = y^2, \quad p(0) = 1, \quad q(0) = 1/2.$$

Note that all the three TPBVP's above satisfy the conditions in step 3.

Next, we combine the solutions of the above TPBVP's with the help of the transformations $t = \varepsilon\tau$ and $t = 1 - \varepsilon\sigma$ in the corresponding domains. The numerical solutions of this example are given in Tables (6.5) - (6.7) along with the asymptotic solutions for different values of ε .

In both of the examples we have taken the tolerence limit $T = 0.0005$ whence the requirement (4.13) is satisfied immediately. However, for specific problems one might need an experienced eye to minimize wasting machine-time in choosing appropriate t_1 and t_2. In the tables the results for the most sensitive state $y(t,\varepsilon)$ out of the trajectories (x,y) is given.

Table 6.2

Numerical Results for Example 3 ($\varepsilon = 0.1$)

t	$y(t,\varepsilon)$		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	1.985000	1.985000	4.098675
0.000001	2.125632	2.210835	4.098545
0.000005	2.228346	2.293701	4.098024
0.00001	2.302382	2.808126	4.097373
0.00005	2.932187	3.123460	4.092168
0.0005	3.201035	3.570831	4.033754
0.001	3.312389	3.705723	3.959164
0.005	3.021392	3.454023	3.454186
0.01	2.8510869	2.851453	3.851472
0.05	-0.972318	-0.975937	-0.975958
0.1	-3.922875	-3.922258	-3.922261
0.2	-6.337021	-6.336842	-6.336836
0.3	-6.638210	-6.631289	-6.631239
0.4	-6.189023	-6.189000	-6.189005
0.5	-5.547271	-5.545323	-5.545291
0.6	-4.893127	-5.894350	-5.894299
0.7	-4.288932	-4.287721	-4.287563
0.8	-3.721013	-3.697003	-3.697209
0.9	-2.923105	-2.998929	-2.998204
0.95	-2.221032	-1.902587	-2.515795
1.0	0.000829	0.000829	-1.845201

Table 6.3

Numerical Results for Example 3 ($\varepsilon = 0.001$)

t	Y (t,ε)		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	1.985000	1.985000	2.007126
0.000001	1.723109	1.823192	1.993972
0.000005	1.689022	1.801236	1.941489
0.00001	1.653125	1.792793	1.876179
0.00005	1.392370	1.369999	1.365296
0.0005	-3.159983	-3.170833	-3.170497
0.001	-6.300897	-6.309312	-6.309256
0.005	-11.013457	-11.012239	-11.012114
0.01	-11.020893	-11.020203	-11.020199
0.05	-10.396324	-10.395291	-10.395398
0.1	-9.662935	-9.660323	-9.660406
0.2	-8.333954	-8.332809	-8.332531
0.3	-7.171729	-7.172530	-7.172223
0.4	-6.157028	-6.156093	-6.156116
0.5	-5.260932	-5.263920	-5.263758
0.6	-4.477029	-4.477192	-4.477197
0.7	-3.781232	-3.780632	-3.780617
0.8	-3.160902	-3.160202	-3.160013
0.9	-2.595230	-2.600001	-2.602911
0.95	-2.000210	-1.078923	-2.344632
1.0	0.000229	0.000222	-2.084680

Table 6.4

Numerical Results for Example 3 ($\varepsilon = 0.00001$)

t	$y(t,\varepsilon)$		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	1.985000	1.985000	1.986211
0.000001	1.000213	0.932124	0.734018
0.000005	-2.990783	-3.002037	-3.191253
0.00001	-5.002317	-5.302938	-3.331528
0.00005	-11.085792	-11.083092	-11.083111
0.0005	-11.166782	-11.164023	-11.164464
0.001	-11.155299	-11.156228	-11.156343
0.005	-11.092005	-11.091487	-11.091569
0.01	-11.011927	-11.011072	-11.011099
0.05	-10.387729	-10.386792	-10.383838
0.1	-9.653991	-9.653029	-9.653043
0.2	-8.326953	-8.326871	-8.326885
0.3	-7.157729	-7.167592	-7.167548
0.4	-6.150809	-6.150999	-6.151807
0.5	-5.255952	-5.259032	-5.259311
0.6	-4.402937	-4.203921	-4.472181
0.7	-3.707621	-3.774029	-3.774646
0.8	-3.000293	-2.892316	-3.152732
0.9	-2.008232	-1.999997	-2.593980
0.95	-1.780236	-1.002138	-2.334749
1.0	0.000002	-1.000000	-2.087062

Table 6.5

Numerical Results for Example 4 ($\varepsilon = 0.1$)

t	$y(t,\varepsilon)$		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	5.984000	5.984000	5.008991
0.000001	5.899231	5.821023	5.008931
0.000005	5.723450	5.692319	5.008692
0.00001	5.529322	5.402132	5.008394
0.00005	5.203124	5.011295	5.006006
0.0005	5.598346	5.492358	5.979208
0.001	5.802341	5.892031	5.949574
0.005	5.690235	5.709235	5.717763
0.01	5.497831	5.448921	5.440750
0.05	3.823092	3.689234	3.659778
0.1	2.272346	2.236002	2.235121
0.2	0.832937	0.844792	0.844804
0.3	0.326821	0.326602	0.326596
0.4	0.124592	0.124492	0.124476
0.5	0.033023	0.033921	0.033984
0.6	-0.018235	-0.019623	-0.019716
0.7	-0.062029	-0.062788	-0.062713
0.8	-0.092135	-0.098899	-0.099565
0.9	-0.113215	-0.114499	-0.114521
0.95	-0.094921	-0.095507	-0.095578
1.0	-0.009009	0.000209	-0.021886

Table 6.6

Numerical Results for Example 4 ($\varepsilon = 0.001$)

t	$y(t,\varepsilon)$		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	5.984000	5.984000	5.984250
0.000001	5.978321	5.978293	5.978269
0.000005	5.958220	5.954889	5.954405
0.00001	5.924002	5.924799	5.924708
0.00005	5.693836	5.692025	5.692407
0.0005	3.633275	3.629008	3.629730
0.001	2.212301	2.202012	2.201641
0.005	0.041927	0.040678	0.040567
0.01	0.000998	0.000402	0.000497
0.05	-0.000987	-0.000682	-0.000625
0.1	-0.003828	-0.003397	-0.003375
0.2	-0.018235	-0.014689	-0.014500
0.3	-0.032729	-0.033295	-0.033125
0.4	-0.059927	-0.059292	-0.059250
0.5	-0.095273	-0.092792	-0.092875
0.6	-0.129985	-0.134995	-0.134000
0.7	-0.183829	-0.182692	-0.182625
0.8	-0.230956	-0.238012	-0.238750
0.9	-0.300125	-0.302257	-0.302375
0.95	-0.123792	-0.321785	-0.337000
1.0	-0.000095	-0.120937	-0.123972

Table 6.7

Numerical Results for Example 4 ($\varepsilon = 0.00001$)

t	$y(t,\varepsilon)$		
	h = 1/50	h = 1/100	Asymptotic solution
0.0	5.984000	5.984000	5.984003
0.000001	5.415832	5.414228	5.414550
0.000005	3.628821	3.629823	3.629482
0.00001	2.202586	2.201872	3.201393
0.00005	0.048201	0.040923	0.040322
0.0005	0.002357	0.000089	0.000002
0.001	0.000092	0.000000	0.000002
0.005	-0.000001	-0.000004	-0.000007
0.01	-0.000098	-0.000032	-0.000035
0.05	-0.002015	-0.000996	-0.000934
0.1	-0.003892	-0.003788	-0.003746
0.2	-0.014026	-0.014823	-0.014995
0.3	-0.033082	-0.033789	-0.033744
0.4	-0.059087	-0.059987	-0.059993
0.5	-0.093000	-0.093788	-0.093741
0.6	-0.138820	-0.134028	-0.134990
0.7	-0.182992	-0.183599	-0.183739
0.8	-0.238892	-0.239888	-0.239988
0.9	-0.303557	-0.303786	-0.303736
0.95	-0.338119	-0.338400	-0.338423
1.0	-0.099267	-0.100278	-0.124989

6.5 Discussions

In § 6.2 - 6.4, we have indicated how the methods developed in the previous chapters might be used for solving optimal control problems numerically. We have concentrated on the special methods suggested in Chapter 4. The Hybrid method is expected to behave well in this application area.

The main feature of the Hamiltonian systems that arise in connection with optimal control problems is that they contain two boundary layers, in general, at the two ends of the interval of integration. The left boundary layer is stable in time whereas the right boundary layer is stable in negative time, the thickness being of $O(|\varepsilon \ln \varepsilon|)$ or roughly $O(\varepsilon)$; (see the assumptions in § 4.2) which we exploit in choosing the stretchings $\tau = (t - t_o)/\varepsilon$ and $\sigma = (t_f - t)/\varepsilon$ at the left and right boundary layers. For the linear problems we assume that the systems are boundary layer observable and boundary layer controlable which guarentee the existence of the boundary layers and in turn, justifies the stretchings. Under these crucial assumptions (see [60-61]), we have obtained solutions to the linear optimal control problems using τ and σ-scaled equations.

The computed results for examples for linear regulators (§ 6.2) show that the solution obtained by the proposed method approximates the original solution better as the mesh is made finer and the parameter ε is made smaller. In the second example, we have taken the reduced solution for comparison purposes, because of the stiffness involved in the problem. Examples are not given for the fixed end point problems and cheap control problems (§ 6.3) for the obvious similarity of these problems with the linear regulators. For non-linear regulators however examples are given to illustrate the methods. We have taken the rapidly varying trajectory component $y(t, \varepsilon)$ for comparison purposes. None of the examples has right boundary-layer behaviour. When we decrease ε to zero, and when the mesh size h is made finer, the convergence of the numerical solution to the reduced one becomes faster. We have given the numerical solutions obtained by the proposed methods along with those by the asymptotic method, in the tables (§ 6.4) hoping that it would facilitate comparison.

We also find that the solutions are close enough except at the boundaries, where the solution

obtained by the proposed method is obviously better than the other. The same dichotomies regarding the stability of the boundary layers still hold for non-linear problems. Our assumptions guarantee that there is no turning point in the middle, whence we are justified in trying to find a solution in the boundary layers with the stretchings τ and σ. Since the thickness of the boundary layers is not yet determined with analytical or asymptotic means for the Hamiltonian systems, and asymptotic methods available to solve this type of problem almost avoid this question, it is suggested that we try to get a solution by the proposed methods where we would be able to know more accurately about the thickness of the boundary layers (rather than just $O(\varepsilon)$ or $O(|\varepsilon \ln\varepsilon|)$. These methods have also the advantage of simplicity over the non-trivial matching procedures for such problems.

Chapter 7

CONCLUSION

The wide applicability of singular perturbation on the one hand and the inherent challenges of singular perturbation on the other stimulated an untiring interest in the mathematicians. It is the latter that attracts every person with a mathematical bent. However, the reason for numerical analysts to poke their noses involves both. A numerical analyst can never be satisfied to know only the nature of a solution of a particular class of problem(s). He is also interested in getting a solution concretely(quantitatively).A primary interest of him is an efficient method. He requires both conceptual and computational efficiency. The method should, according to a numerical analyst, be simple, efficient and accurate. The first and foremost criterion is accuracy. To be a accurate, a numerical method must be convergent (i.e., non-accumulation of errors and etc.) and the numerical solution that it produces must possess all the qualitative properties.

In trying to find out numerical methods for singularly perturbed two-point boundary-value problems, we have implicitly narrowed down our interest. This limitation of scopes of the applicability of the methods was guided by two essential properties : the boundary-layer phenomena and oscillation. We have considered problems whose solutions show boundary layer behaviours by excluding the possibility that the eigen-values of the system matrix be purely imaginary and eradicating oscillations by smoothness conditions. Singularly perturbed systems having boundary layer behaviours always possess two modes : slow state and fast state (in control theory language.). The fast state again consists of two modes : rapidly growing and rapidly decaying. These properties are projected well in decoupling the slow and fast modes in Chapter 2. The decoupling transformation introduced has the advantage of simplicity over the usual Lypunov-type transformation The analytic decoupling (compared to the decoupling of difference schemes) facilitates applying Euler's schemes to the fast mode. However, a similar problem remains there such as decoupling the difference equations (by an implicit multi-step method) by

using Lypunov-type transformations [62] (parallel to Matheij [74]). We emphasize that an implicit scheme rather than an explicit scheme is to be used, due to the failure of Lypunov-type transformations on the difference equations yielded by the explicit one-step method. (see § 2.4 - 2.5).

Guided by simplicity and efficiency, we have also suggested some special methods to handle some sub-classes of the singular perturbation problems, namely, the problems without turning points. Though we have only seen those special methods to work well for homogeneous problems, they are expected to work well also for non-homogeneous ones. For non-linear problems, the special methods have the advantage that they can be applied with the usual process of linearization. In contrast the general purpose method, we have developed in Chapter 2 needs the linearization of a non-linear problem before applying the difference schemes. This was achieved by approximating the non-linear function $F(t,z)$, (in $\dot{z} = F(t,z)$) by a sequence of linear functions (see Chapter 5).

The numerical experimentations carried out for the reliability and realizability of the methods

developed report on the affirmative. It was observed that the methods approximate the exact/ asymptotic solutions with less labour. When the parameter ε approaches zero, the numerical solution also approaches the reduced solution (preserving the non-uniformity characteristic of such problems). It was also observed that the expected accuracy could always be achieved without much effort.

There are at least two ways to combat stiffness involved in the computation of solutions of singularly perturbed systems. One is to design a better computer (hardware) and the other is to design a better algorithm (soft ware). These two are not necessarily mutually exclusive. However, our emphasis was on the latter alternative. ...

All the computations were carried out on DEC-1090 computer system available at Indian Institute of Technology, Kanpur.

List of References

1. Abrahamsson, L.R., H.B. Kreiss and H.O. Kreiss (1974), Difference approximations for singular perturbations of systems of ordinary differential equations; Num.. Math., 22, 367.

2. Aiken, R.C. (1985), Ed.,Stiff Computation, Oxford University Press, New York, Oxford.

3. Aiken, R.C. and L. Lapidus (1975), Problem approximation for stiff ordinary differential equations; AICHE Journal, 21, 627.

4. Ardema, M.D. (1983), An introduction to singular perturbations in nonlinear optimal control; p. 1 in [5].

5. Ardema, M.D. (1983), Ed., Singular Perturbations in Systems and Control, CISM courses and lectures no. 280, International Center for Mechanical Sciences, Springer Verlag, New York.

6. Ascher, U. (1980), Solving boundary-value problems with a spline-collocation code; J. Comput. Phys., 34, 401.

7. Ascher, U. (1985), On some difference schemes for singular singularly perturbed boundary-value problems; Num. Math., 46, 1.

8. Ascher, U. and R. Weiss (1983), Collocation for singular perturbation problems I : first order systems with constant coefficients; SIAM J. Num. Anal., 20, 537.

9. Ascher, U. and R. Weiss (1984), Collocation for singular perturbation problems II : linear first order systems without turning points; Math. Comp., 43, 157.

10. Ascher, U. and R. Weiss (1984), Collocation for singular perturbation problems III : nonlinear problems without turning points; SIAM J. Sci. Stat. Comp., 5, 811.

(In case of journal articles, the first number is the volume no. and the last is the beginning page number.)

11. Ascher, U., J. Christiansen and R.D. Russell (1979),COLSYS - a collocation code for boundary value problems; appeared in [18].

12. Axelsson, O. and G.F. Carey (1985), On the numerical solution of two-point singularly perturbed boundary value problems; Comp. Methods Appl. Mech. Engg., 50, 217.

13. Bellman, R. (1964), B., Perturbation techniques in mathematics, physics and engineering, Holt, Rinehart, Winston, New York.

14. Benoit, E., J.L. Collot, F. Diener et M. Diener (1980), Chasse all canard; Publication IRMA, Strasbourg.

15. Berger, A.E., J.M. Solomon and M. Ciment (1981), An analysis of a uniformly accurate difference method for a singular perturbation problem; Math. Comp., 37, 79.

16. Bucy, R.S. (1967), Two-point boundary-value problems of linear Hamiltonian systems; SIAM J. Appl. Math., 15, 1385.

17. Campbell, S.L. and N.J. Rose (1979), Singular perturbation of autonomous linear systems; SIAM J. Math. Anal. 10, 542.

18. Childs, B., M. Scott, J.W. Daniel, E. Denman and P. Nelson (1979), Ed., Codes for boundary-value problems in ordinary differential equations, Lecture Notes in Computer Science No. 76, Springer-Verlag.

19. Cohen, D.S. (1973), Singular perturbation of nonlinear two-point boundary-value problems; J. Math. Anal. Appl., 43, 151.

20. Coppel, W.A. (1965), B., Stability and asymptotic behaviour of differential equations, D.C. Heath, Boston.

21. Coppel, W.A. (1967), Dichotomies and reducibility, J. Diff. Eqns., 3, 500.

22. Dekker, K. and J.W. Verwer (1984), B., Stability of Runge-Kutta methods for stiff nonlinear differential equations, North-Holland, Amsterdam, New York, Oxford.

23. Dontchev, A.L. (1983), B., Perturbations, approximations and sensitivity analysis of optimal control systems, Lecture Notes in Control and Information Sciences, No. 52, Springer-Verlag, New York.

24. Dorr, F.W. (1970), The numerical solution of singular perturbations of boundary-value problems; SIAM J. Num. Anal., 7, 281.

25. Dorr, F.W. and S.V. Parter (1970), Singular perturbations of nonlinear boundary-value problems with turning points; J. Math. Anal. Appl.,29, 273.

26. Dorr, F.W., S.V. Parter and L.F. Shampine (1973), Applications of the maximum principle to singular perturbation problems; SIAM Review, 15, 43.

27. Eckhaus, W. (1973), B., Matched asymptotic expansions and singular perturbations, North-Holland Publ. Co., Amsterdam.

28. Eckhaus, W. (1979), B., Asymptotic analysis of singular perturbations, North-Holland Publ. Co., Amsterdam.

29. Engstorm, P.G. (1982), Singular perturbations of two Cauchy problems in a Hilbert space; Rocky Mount. J. Math., 12, 561.

30. Erdelyi, A. (1956), B., Asymptotic expansions, Dover, New York.

31. Ferguson, W.E., Jr. (1975), A singularly perturbed linear two-point boundary-value problem; Ph.D. Thesis, California Institute of Technology, Pasadena, and University Mico-films.

32. Ferguson, W.E., Jr. (1986), Analysis of a singularly perturbed linear two-point boundary-value problem; SIAM J. Num. Anal., 23, 940.

33. Finden, W.F. (1983), An asymptotic approximation for singular perturbations; SIAM J. Appl. Math., 43, 107.

34. Flaherty, J.E. and W. Mathon (1980), Collocation with polynomial and tension splines for singularly perturbed boundary-value problems; SIAM J. Sci. Stat. Comput., 1, 260.

35. Flaherty, J.E. and R.E. O'Malley Jr. (1977), The numerical solution of boundary-value problems for stiff differential equations; Math. Comp., 31, 66.

36. Freedman, M.I, and B. Granoff (1976), Formal asymptotic solutions of a singularly perturbed nonlinear optimal control problem; J. Optim. Theo. Appl., 19, 301.

37. Freedman, M.I. and J.L. Kaplan (1976), Singular perturbations of two-point boundary-value problems arising in optimal control; SIAM J. Control Optim., 14, 789.

38. Gingold, H. (1978), A method of global block-diagonalization for matrix valued functions; SIAM J. Math. Anal., 9, 1076.

39. Habets, P. (1985), Total and robust asymptotic stability in singular perturbation problems; In : Large scale systems, Decision Making, Mathematics and Control; A bridge between control science and technology, Ninth Triennial World Congress of IFAC.

40. Hadlock, C.A. (1973), Existence and dependence on a parameter of solutions of a nonlinear two-point boundary-value problem; J. Diff. Eqns., 14, 498.

41. Harris, W.A., Jr. (1973), Singularly perturbed boundary-value problems revisited; In : Lecture Notes in Math., V. 312, p. 54, Springer-Verlag, New York.

42. Hemker, P.W. (1977), A numerical study of stiff two-point boundary problems; MCT 80, Mathematical Centre, Amsterdam.

43. Hemker, P.W. and J.J.H. Miller (1979), Ed., Numerical analysis of singular perturbation problems, Academic Press, New York.

44. Hoppensteadt, F.C. (1971), Properties of solutions of ordinary differential equations with small parameters; Commun. Pure and Appl. Math., 24, 807.

45. Howes, F.A. (1984), Boundary layer behaviour in perturbed second order systems; J. Math. Anal. Appl., 104, 467.

46. Hsiao, G.C. and K.E. Jordan (1979), Solutions to the difference equations of singular perturbation problems; In : [43], p. 433.

47. Kadalbajoo, M.K. and Y.N. Reddy (1986), A new approach for solving singular perturbation problems; J. Comp. Phy., 62, 349.

48. Kadalbajoo, M.K. and Y.N. Reddy (1987), An initial value technique for a class of nonlinear singular perturbation problems; J. Optim. Theo. Appl., 53, 395.

49. Kadalbajoo, M.K. and Y.N. Reddy (1987), Numerical treatment of singularly perturbed two-point boundary-value problems ; Appl. Math. Comp., 21, 93.

50. Kadalbajoo, M.K. and Y.N. Reddy (1987), Approximate method for the numerical solution of singular perturbation problems; Appl. Math. Comp., 21, 185.

51. Kadalbajoo, M.K. and Y.N. Reddy (1987), A non-asymptotic method for general linear singular perturbation problems; J. Optim. Theo. Appl., 55, 257.

52. Kalman, R.E. (1960), Contributions to the theory of optimal control; Bol. Soc. Mat. Mexicana, 5, 102.

53. Kato, T. (1966), B., Perturbation theory of linear operators, Springer-Verlag, Berlin.

54. Kato, T. (1982), B., A short introduction to perturbation theory for linear operators, Springer-Verlag, New York.

55. Keller, H.B. (1968), B., Numerical methods for two-point boundary-value problems, Blaisdell.

56. Keller, H.B. and A.B. White Jr. (1975), Difference methods for boundary-value problems in ordinary differential equations; SIAM J. Num. Anal., 12, 791.

57. Kellog, R.B. and A. Tsan (1978), Analysis of some difference approximations for singular perturbation problems without turning points; Math. Comp., 32, 1025.

58. Kevorkian, J. and J.D. Cole (1981), B., Perturbation methods in applied mathematics, Springer-Verlag, New York.

59. Khalil, H.K. (1981), Asymptotic stability of nonlinear multiparameter singularly perturbed systems; Automatica, 17, 797.

60. Kokotovic, P.V. (1984), Applications of singular perturbation techniques to control problems; SIAM Review, 26, 501.

61. Kokotovic, P.V. and R.A. Yackel (1972), Singular perturbation of linear regulators : basic theorems; IEEE Trans. on Aut. Cont., AC - 17, 29.

62. Kokotovic, P.V., H.K. Khalil and J. O'Reilly (1986), B., Singular perturbation methods in control : analysis and design, Academic Press, New York.

63. Kokotovic, P.V., R.E. O'Malley Jr. and P. Sannuti (1976), Singular perturbations and order reduction in control theory - an overview; Automatica, 12, 123.

64. Kopell, N. and S.V. Parter (1981), A complete analysis of a model non-linear singular perturbation problem having a continuous locus of singular points; Adv. Appl. Math., 2, 212.

65. Kreiss, H.O. (1972), Centered difference approximation to singular systems of ODEs; Symposium Mathematica X, Instituto Nazionale di Alta Mathematica.

66. Kreiss, H.O. (1976), B., Numerical methods for singular perturbation problems;SIAM-AMS proceedings, Vol. - 10, American Mathematical Society, Providence RI.

67. Kreiss, B. and H.O. Kreiss (1981), Numerical methods for singular perturbation problems; SIAM J. Num. Anal., 18, 262.

68. Kreiss, H.O., N.K. Nichols and D.L. Brown (1986), Numerical methods for stiff two-point boundary-value problems; SIAM J. Num. Anal., 23, 325.

69. Lang, S. (1983), B., Real analysis, Addison-Wesley.

70. Linz, P. (1979), B., Theoretical numerical analysis (an introduction to advanced techniques), John Wiley and Sons, New York.

71. Lord Rayleigh and J.W. Strutt (1945), B., The theory of sound, Vol.-1, 2nd Ed., Dover, New York.

72. Lorenz, J. (1979), Combinations of initial and boundary-value methods for a class of singular perturbation problems; In [43].

73. Lutz, R. and M. Goze (1981), B., Nonstandard analysis : a practical guide with applications, Lecture Notes in Mathematics, Vol.-881, Springer-Verlag, New York.

74. Matheij , R.M.M. (1979), On approximating smooth solutions of linear singularly perturbed ODE ; In [43], p. 457.

75. Miranker, W.L. and J.P. Morreeuw (1974), Semi-analytic numerical studies of turning points arising in stiff boundary-value problems; Math. Comp., 28, 1017.

76. Nayfeh, A.H. (1979), B., Perturbation methods, Wiley, New York.

77. Nayfeh, A.H. (1981), B., Introduction to perturbation techniques, Wiley, New York.

78. Niederdrenk, K. and H. Yserentant (1983), Die gleichmäßige stabilität singular gestörter diskreter und kontinuierlicher randwert probleme; Num. Math., 41, 223.

79. Nipp, K. (1983), An extension of Tikhonov's theorem in singular perturbations for the planar case; ZAMP, 34, 277.

80. Nørseett, S.P. and G. Wanner (1981), Perturbed collocation and Runge-Kutta methods; Num. Math., 38, 193.

81. O'Malley, R.E., Jr. (1972), The singularly perturbed linear state regulator problem; SIAM J. Control, 10, 399.

82. O'Malley, R.E., Jr. (1974), B., Introduction to singular perturbations, Academic Press, New York.

83. O'Malley, R.E., Jr. (1978), Singular perturbation and optimal control; In : Mathematical Control Theory, Lecture Notes in Mathematics, Vol.-680, p. 170, Springer-Verlag, New York.

84. O'Malley, R.E., Jr. (1979), A singular singularly perturbed linear boundary-value problem; SIAM J. Math. Anal., 10, 695.

85. O'Malley, R.E., Jr. (1982), Book reviews; Bulletin (New series) of the American Math. Soc., 7, 414.

86. O'Malley, R.E., Jr. and L.R. Anderson (1979), Singular perturbation, order reduction and decoupling large-scale systems; In [43], p. 317.

87. O'Malley, R.E., Jr. and A. Jameson (1975), Singular perturbations and singular arcs - part I; IEEE Trans. Aut. Control, AC-20, 218.

88. O'Malley, R.E., Jr. and A. Jameson (1977), Singular perturbations and singular arcs - part II; IEEE Trans. Aut. Control, AC-22, 328.

89. Ortega, J.M. and W.C. Rheinboldt (1970), B., Iterative solution of nonlinear equations in several variables, Academic Press, New York.

90. Ortiz, E.L. (1984), An error analysis of the Tau method for a class of singularly perturbed problems for differential equations; Math. Methods in Appl. Sci., 6, 457.

91. Osher, S. (1980), Nonlinear singular perturbation problems and one-sided difference schemes; SIAM J. Num. Anal., 18, 129.

92. Pearson, C.E. (1968), On a differential equation of boundary layer type; J. Math. and Phy., 47, 134.

93. Pearson, C.E. (1968), On a nonlinear differential equation of boundary layer type; J. Math. and Phy., 47, 351.

94. Petrovskii, I.G. (1966), B., Ordinary differential equations, Rev. Eng. Ed., Englewood Cliffs, N.J., Prentice Hall.

95. Pontryagin, L.S., V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mischenko (1962), B., The mathematical theory of optimal process, Interscience, New York.

96. Prandtl, L. (1905), Über flüssigkeit-bewegung bei kleiner reibung; Verh. III Inter. Math. Kongresses, Tuebner, Leipzig, p. 484.

97. Roberts, S.M. (1982), A boundary-value technique for singular perturbation problems; J. Math. Anal. Appl., 87, 489.

98. Roberts, S.M. (1983), The analytical and approximate solutions of $\varepsilon y'' = yy'$; J. Math. Anal. Appl., 97, 245.

99. Roberts, S.M. (1984), Solution of $\varepsilon y''+yy'-y = 0$ by a non-asymptotic method; J. Optim. Theo. Appl., 44, 303.

100. Sakai, M. (1975), Approximate method for the integration of singular perturbation problems; Mem. Fac. Sci., Kyushu Univ., 29, 185.

101. Sakai, M. and R.A. Usman (1984), A posteriori improvement of cubic-spline approximate solutions of two-point boundary-value problems, Congressus Numerantium, 42, 265.

102. Sannuti, P. (1974), Asymptotic series solution of singularly perturbed optimal control problems; Automatica, 10, 183.

103. Sannuti, P. (1975), Asymptotic expansions of singularly perturbed quasilinear optimal systems; SIAM J. Control, 13, 572.

104. Smith, D.R. (1980), On a singularly perturbed boundary-value problem arising in the physical theory of semiconductors; Tech. Rep. TUM - M8021, Technische Universität Miinchen.

105. Tikhonov, A.N. (1952), Systems of differential equations containing a small parameter multiplying the derivative; Mat. Sb., 31,(73),575 (In Russian).

106. Tikhonov, A.N., A.B. Vasileva and A.G. Sveshnikov (1985), B., Differential equations, Springer-Verlag, New York.

107. van Dyke, M. (1964), B., Perturbation methods in fluid dynamics, Academic Press, New York.

108. Vasileva, A.B. (1963), Asymptotic behaviour of solutions to certain problems involving nonlinear differential equations containing a small parameter multiplying the highest derivatives; Russ. Math. Sur., 18, 13.

109. Vasileva, A.B. and V.F. Butuzov (1973), B., Asymptotic expansions of solutions of singularly perturbed differential equations, Nauka, Moscow (In Russian).

110. Vasileva, A.B. and V.F. Butuzov (1978), B., Singularly perturbed systems in critical cases, Moscow University Press, Moscow (In Russian).

111. Vasileva, A.B. and V.M. Volosov (1967), The work of Tikhonov and his approach and his pupils on ordinary differential equations containing a small parameter; Russ. Math. Sur., 22, 124.

112. Wasow, W. (1965), B., Asymptotic expansions for ordinary differential equations, Interscience, New York.

113. Weiss, R. (1984), An analysis of the box and trapezoidal schemes for linear singularly perturbed boundary-value problems; Math. Comp., 42, 41.

114. Wilde, R.R. and P.V. Kokotovic (1973), Optimal open and closed loop control of singularly perturbed linear systems; IEEE Trans. Aut. Control., AC-18, 616.

115. Wilkinson, J.H. (1965), B., The algebraic eigenvalue problem, Clarendon Press, Oxford.

116. Wonham, W.M. (1979), B., Linear multivariable control : a geometric approach, Springer-Verlag, New York.

Date Slip